ANL/EVS-17/9

Climate Change Impacts at Department of Defense Installations

Environmental Science Division

About Argonne National Laboratory
Argonne is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC under contract DE-AC02-06CH11357. The Laboratory's main facility is outside Chicago, at 9700 South Cass Avenue, Argonne, Illinois 60439. For information about Argonne and its pioneering science and technology programs, see www.anl.gov.

DOCUMENT AVAILABILITY

Online Access: U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via DOE's SciTech Connect (http://www.osti.gov/scitech/)

Reports not in digital format may be purchased by the public from the National Technical Information Service (NTIS):

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312
www.ntis.gov
Phone: (800) 553-NTIS (6847) or (703) 605-6000
Fax: (703) 605-6900
Email: **orders@ntis.gov**

Reports not in digital format are available to DOE and DOE contractors from the Office of Scientific and Technical Information (OSTI):

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
www.osti.gov
Phone: (865) 576-8401
Fax: (865) 576-5728
Email: **reports@osti.gov**

ANL/EVS-17/9

Climate Change Impacts at Department of Defense Installations

prepared by
Rao Kotamarthi, Environmental Science Division, Argonne National Laboratory
Jiali Wang, Environmental Science Division, Argonne National Laboratory
Zach Zoebel, Department of Atmospheric Sciences, University of Illinois, Urbana, IL
Don Wuebbles, Department of Atmospheric Sciences, University of Illinois, Urbana, IL
Katharine Hayhoe, Climate Science Center, Texas Tech, Lubbock, TX
Michael Stein, Department of Statistics, University of Chicago, Chicago, IL
David Changnon, Department of Geography, Northern Illinois University, DeKalb, IL

Prepared for the U.S. Department of Defense, Strategic Environmental Research and Development Program

June 16, 2017

This page intentionally left blank.

CONTENTS

FIGURES

FIGURES (Cont.)

FIGURES (Cont.)

FIGURES (Cont.)

TABLES

LIST OF ACRONYMS

ALCF	Argonne Leadership Computing Facility
ARRM	Asynchronous Regional Regression Model
CAM	Community Atmosphere Model
CCSM	Community Climate System Model
CESM	Coupled Earth System Model
CMIP5	Coupled Model Intercomparison Project 5
CONUS	continental United States
CPC	Climate Prediction Center
CRU	Climatic Research Unit
DFT	discrete Fourier transform
ETCCDI	Expert Team on Climate Change Detection and Indices
GCM	general circulation model
GEV	generalized extreme value
LBC	lateral boundary condition
NAM	North American Monsoon
NARCCAP	North American Regional Climate Change Assessment Program
NARR	North American Regional Reanalysis
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCEP-R2	National Centers for Environmental Prediction—U.S. Department of Energy Reanalysis II
NEON	National Ecological Observatory Network
NERSC	National Energy Research Scientific Computing Center
NOAA	National Oceanic and Atmospheric Administration
PDF	probability density function
PRISM	Precipitation-Elevation Regressions on Independent Slopes Model
RCM	regional climate model
RCP	Representative Concentration Pathway
REA	reliability ensemble averaging
RMSE	root-mean-square error
SD	standard deviation
SDM	statistical downscaling model
TCC	temporal correlation coefficient
TRMM	Tropical Rainfall Measuring Mission
UDEL	University of Delaware
WRF	Weather Research and Forecasting
WRFG	Weather Research and Forecasting with Grell-Devenyi convective parameterization
WSM6	WRF Single-Moment 6-Class

KEYWORDS

Climate Downscaling, Dynamic Downscaling, Statistical Downscaling, Climate Projections, DoD Installations

ACKNOWLEDGMENTS

Funding for this research provided by the Strategic Environmental Research and Development Program (SERDP) under project RC-2242. This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This research used resources of the Argonne Leadership Computing Facility at Argonne National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-06CH11357.

This page intentionally left blank.

ABSTRACT

a. Objective: This project is aimed at providing the U.S. Department of Defense (DoD) with a comprehensive analysis of the uncertainty associated with generating climate projections at the regional scale that can be used by stakeholders and decision makers to quantify and plan for the impacts of future climate change at specific locations. The merits and limitations of commonly used downscaling models, ranging from simple to complex, are compared, and their appropriateness for application at installation scales is evaluated. Downscaled climate projections are generated at selected DoD installations using dynamic and statistical methods with an emphasis on generating probability distributions of climate variables and their associated uncertainties. The sites selection and selection of variables and parameters for downscaling was based on a comprehensive understanding of the current and projected roles that weather and climate play in operating, maintaining, and planning DoD facilities and installations.

b. Technical Approach: Using interviews with DoD installation stakeholders, we have identified climate variables and key vulnerabilities of interest in the initial phase of our project. We generated high-resolution climate projections for the North American continent by combining state-of-the-art dynamical and statistical downscaling models with quality-controlled observations and the latest simulations from global models of the Earth's climate system. Expected changes in climate variables were evaluated by analyzing these downscaled climate model products at the relevant spatial scales for selected DoD built installations.

c. Benefits: This project generated the knowledge base, data sets, and tools needed for making preliminary assessments of vulnerabilities due to climate change (e.g., severe-event probabilities), expected changes in operating parameters such as heating and cooling needs, and other potential challenges for DoD installations and range management. Specifically, by evaluating the value added and the appropriateness of downscaled projections to the size of a DoD installations, we significantly advance the appropriate use of downscaled climate projections. By generating downscaled projections and the associated uncertainty at specific locations, we provided an important resource for DoD to use in planning to adjust to changing local environments that may affect DoD facilities and installations. The methodology, observational data sets, and model products produced during the course of the project are available as an easily accessible database for future use by DoD. An illustrative example developed will be made available in a separate document for DoD use. A GIS application for further exploration of the data at selected installations will also be provided. Together, these products will meet the ultimate goal of assisting DoD in developing informed policies when confronting future change.

1 OBJECTIVE

Typically, global climate change projections do not include a clearly defined analysis of the associated uncertainties at small spatial scales—information that is essential for the U.S. Department of Defense's (DoD's) needs. Therefore, we are estimating the uncertainties in the downscaled projections associated with climate change for DoD installations. This is one of the key desired outcomes of the State of Need (SON). Specifically, we conducted new simulations with one of the best regional climate models (RCMs) for simulating climate over the United States—the Weather Research and Forecasting (WRF) model—to develop a dataset for estimating model bias when simulating historical climate data and account for uncertainties in projections of future climate. Our focus is on generating as many ensemble members as possible with WRF (different initial and boundary conditions, as well as different physics parameters). Statistically downscaled estimates of climate projections from multi-model Coupled Model Intercomparison Project 5 (CMIP5) global climate ensembles also contribute to this uncertainty analysis.

The primary technical objectives of our SERDP project are as follows:

- Identify the role of weather information at DoD installations; develop potential-use cases and data needs for climate projections.
- Develop an understanding of the views of stakeholders on climate- and weather-related questions through interviews and questionnaires.
- Generate downscaled climate projections and associated uncertainties for DoD installations.
- Develop an illustrative example that uses high-resolution projections and associated uncertainties for informing DoD infrastructure planning and management.

As the initial step in this project, we collected and evaluated available weather and climate data for specific DoD facilities of interest and established direct contact with DoD installation staff to document their use of day-to-day weather and, more generally, climate information. Interactions with DoD and associated facility stakeholders were helpful in selecting the key vulnerabilities and downscaled variables for our analyses. In the second step, we developed relevant climate projections at the local scale for selected DoD facilities. The results of this effort are published in a wide variety of peer-reviewed journal articles and a summary is provided of the results is provided in Section 4.

2 BACKGROUND

Expectations and projections of climate change are based on large numerical models of the Earth system that simulate physical processes in atmospheric, oceanic, biospheric, and landscape systems. Such models have made rapid progress in simulating the climate observed over the past few decades. However, because of the enormous complexity of the processes that need to be modeled and the spatial scales and time scales required to simulate climate scenarios, a large amount of resources is necessary to develop and perform global- and regional-scale simulations. Computational limitations relate to the ability to represent small-scale processes (spatially and temporally) in general circulation models (GCMs), which many currently available models can resolve only on the order of 100 km or more.

The most recent version of the National Center for Atmospheric Research (NCAR) Coupled Earth System Model (CESM) 1.0 has a spatial resolution of 50 km. However, applying CESM at this spatial resolution over multiple decades requires computing resources beyond those of all but the largest high-performance computing clusters. In contrast, climate change impacts relevant to stakeholders occur at smaller scales—often, they affect specific facilities or other specific locations. Vulnerability assessments and resource planning are limited by the present inability of GCM projections to provide model results at these spatial and time scales. To overcome this limitation in the GCM scale resolution, a broad suite of dynamical and statistical methods—collectively known as downscaling models—have been developed to provide a reasonable method for generating climate model output at the scale required to respond to the urgent needs of decision makers.

To develop a consistent set of climate projections for specific DoD installations, we combined state-of-the-art, well-evaluated statistical methods with regional climate simulations performed for this project at a spatial resolution of 12 km and the latest products available from CMIP5. The site-specific climate projections will be presented for the historical period (1995–2004), the mid-21st century (2045–2055), and the end of the 21st century (2085–2095) for further analysis using an GIS framework by early 2017.

The technical approach used in the proposal is shown in Figure 1. The downscaling activity to generate the necessary model output for estimating climate changes at selected DoD installations uses both statistical and dynamic downscaling with uncertainty estimation methods, as described below.

2.1 STATISTICAL DOWNSCALING

Statistical downscaling produces high-resolution projections at the scale of the observations (individual weather stations or gridded observations) from global model output by developing relationships between historical global model simulations and observed conditions at the location(s) of interest. This method is very flexible; it can be tuned to obtain finer-resolution output for targeted variables and for selected locations. The ease of use of this method and its flexibility have led to a wide variety of applications for assessing impacts of climate change

(e.g., Kattenberg et al. 1996; Hewitson and Crane 1996; Giorgi et al. 2001; Wilby et al. 2004; and references therein). Approaches encompass a range of statistical techniques, from simple linear regression to more complex applications based on weather generators (Wilks and Wilby 1999), canonical correlation analysis (e.g., von Storch et al. 1993), or artificial neural networks (e.g., Crane and Hewitson 1998). Team members have used more recent versions of these statistical methods to provide the basis for regional climate assessments for various states, regions, and government agencies (e.g., Hayhoe et al. 2004, 2008, 2010a; USGCRP 2009).

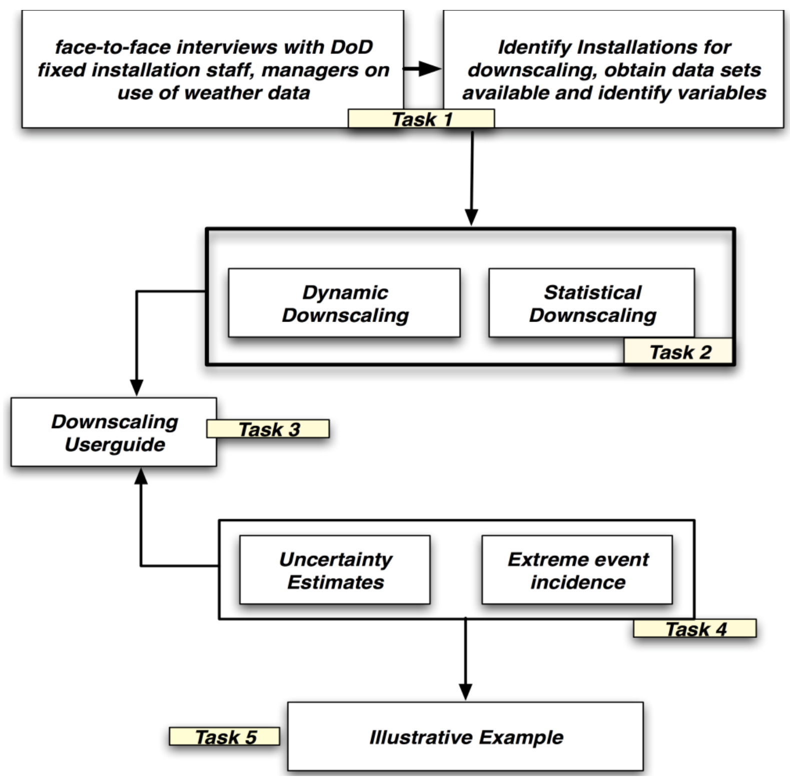

FIGURE 1 Technical Approach and Specific Tasks Identified

2.2 DYNAMICAL DOWNSCALING

Dynamical downscaling by RCMs generally refers to the use of limited-area climate models that are forced at the boundaries by output from a GCM or a relatively large-scale reanalysis dataset (Giorgi and Mearns 1991, 1999; McGregor 1997; Wang et al. 2004). Regional models were developed primarily by adapting mesoscale meteorological forecasting models to climate scales. As a result, most RCMs (including the two models used in this project) include detailed representations of land surface process, cloud physics, and radiative transfer. Diagnostic simulations using RCMs are usually based on reanalysis datasets available from the National Centers for Environmental Prediction (NCEP; e.g., Kanamitsu et al. 2002) or similar sources (ERA-40 developed by the European Centre for Medium Range Weather Forecasts; Uppala et al. 2005). Typically, RCMs operate at spatial scales ranging from 10 to 50 km. Very high spatial resolution, of the order of a few kilometers, can be achieved by using models with even more

limited areas (e.g., Grell et al. 2000). Higher spatial resolution generally improves a model's ability to simulate spatially variable fields such as precipitation, temperature, and atmospheric circulation; thus, in many instances RCMs can produce a more accurate forecast (especially for extreme events) at regional scales. These models have been used widely in applications requiring regional resolution, particularly when higher-resolution climate projections are needed to estimate potential climate impacts on air quality and hydrology.

2.3 ESTIMATION OF UNCERTAINTY

Hawkins and Sutton (2009) ascribed uncertainties in regional-scale climate projections to three primary causes: (1) model internal variability; (2) model response uncertainty or the climate forcing sensitivity of the model, which includes structural and parametric uncertainty; and (3) scenario uncertainty. Model uncertainty and internal variability are estimated to account for 50–80% of the uncertainty in regional forecasts for the next several decades, while scenario uncertainties become dominant in temperature projections toward the end of the century. To resolve the range of uncertainty due to model response in weather forecasting, multi-model ensembles have been applied (Krishnamurti et al. 2000; Palmer 2000). The climate modeling community has adopted this multi-model ensemble approach to evaluate the robustness of projections of future climate change under various forcing scenarios, to process parameter choices, to model structure, and to express the resulting uncertainty.

A good example of this kind of approach is the reliability ensemble averaging (REA) method proposed by Giorgi and Mearns (2002). This method assigns reliability classifications for the multi-model ensemble simulation by using two metrics that evaluate model performance and model convergence. Model performance is determined from differences or bias between observational data and model hindcasts. The convergence criterion uses the distance of the model forecast from the final multi-model consensus over the future projection. Tebaldi et al. (2004, 2005) and Smith et al. (2009) have developed Bayesian analysis methods to quantify uncertainty, for example where hierarchical Bayesian models incorporate the criteria of the REA method in a formal probabilistic framework to derive the probability density function (PDF) of present and future temperature and its change at the regional level. Recent analyses have discouraged the use of Bayesian likelihoods on the basis of model biases; instead, we use model weighting based on the GCM's ability to simulate large-scale atmospheric dynamics that are relevant to the continental United States, as derived from the literature (e.g., Stoner et al. 2009, 2012) and generate an ensemble using three different GCMs to provide the boundary conditions for the downscaling calculations discussed here.

A significant task in uncertainty classification is estimating the model bias with high-quality, high-spatial-resolution observational datasets. We use an extensive model bias evaluation to quantify the uncertainty with the methods described above. Since we are conducting simulations for only two Representative Concentration Pathway (RCP) scenarios and have only a few ensemble members, we cannot analyze the scenario uncertainty; however, when combined with the models used in the North American Regional Climate Change Assessment Program (NARCCAP) project, one could provide weighted-mean uncertainty. We have also tested the uncertainty due to parameter choice and evaluated the model performance and

sensitivity for a few physics parameters (e.g., cumulus parameterization and microphysics scheme) identified during model evaluation, and we are generating an estimate of model-observational bias. In addition, we conducted a larger number of short-time-scale simulations to estimate the model internal variability as a result of using different boundary conditions and initializations (see Section 4.4.2).

3 TECHNICAL APPROACH

The project used a detailed technical approach to address the issues discussed in the background section. Below is a description of the materials and methods used in the research and the simulations performed.

3.1 IDENTIFICATION OF INSTALLATIONS FOR DOWNSCALING

We conducted a survey of available meteorological data and climate diversity to identify 12 DoD installations for climate downscaling, producing downscaled climate projections for the entire domain shown in Figure 2a. The selected installations are distributed throughout the continental United States (CONUS) and Alaska (Figure 2b). For these locations, we will provide GID accessible files for analysis and plotting as a part of this project by early 2017.

3.2 USE OF WEATHER DATA AND VIEWS ON CLIMATE CHANGE AMONG THE DOD SITE CONTACTS

To identify the weather and climate change information uses and needs of DoD installation decision makers, we disseminated a questionnaire to DoD stakeholders at 10 U.S. installations (8 Army installations, 1 Air Force installation, and 1 Marine installation). This process was facilitated through DoD liaisons working with SERDP Climate Change projects. The 34 questionnaires completed and returned considered the use of weather and historical climate data to guide current decisions and the use of climate change projections in future endeavors.

Weather directly affected 33 of 34 stakeholders. Weather information and short-term forecasts appear to be used extensively in daily to weekly decision making at all installations. Uses varied among the stakeholders in environmental, sustainability management, conservation, operations and management, emergency managers, and master planning positions. However, when asked to identify weather extremes that directly affected their activities and decisions, stakeholders indicated that heavy, short-duration rainfall-associated flooding events and drought-associated heat waves caused the greatest number of installation impacts. When asked whether they used historical climate information to determine how frequently these extremes occur, most said “no,” while others provided anecdotal information.

Climate change estimates and model projections were not provided to stakeholders or are not being used in current decision making efforts by 31 of 34 participants. The primary reasons for not using the projections are related to the specific missions of the stakeholder group. Most (19 of 34) indicated that if climate change estimates were made available, they would not incorporate that information into current or future decisions. For those who could see potential uses for such projections, most wanted future projections of precipitation that could lead to too wet or too dry conditions. Most stakeholders were not comfortable addressing issues related to

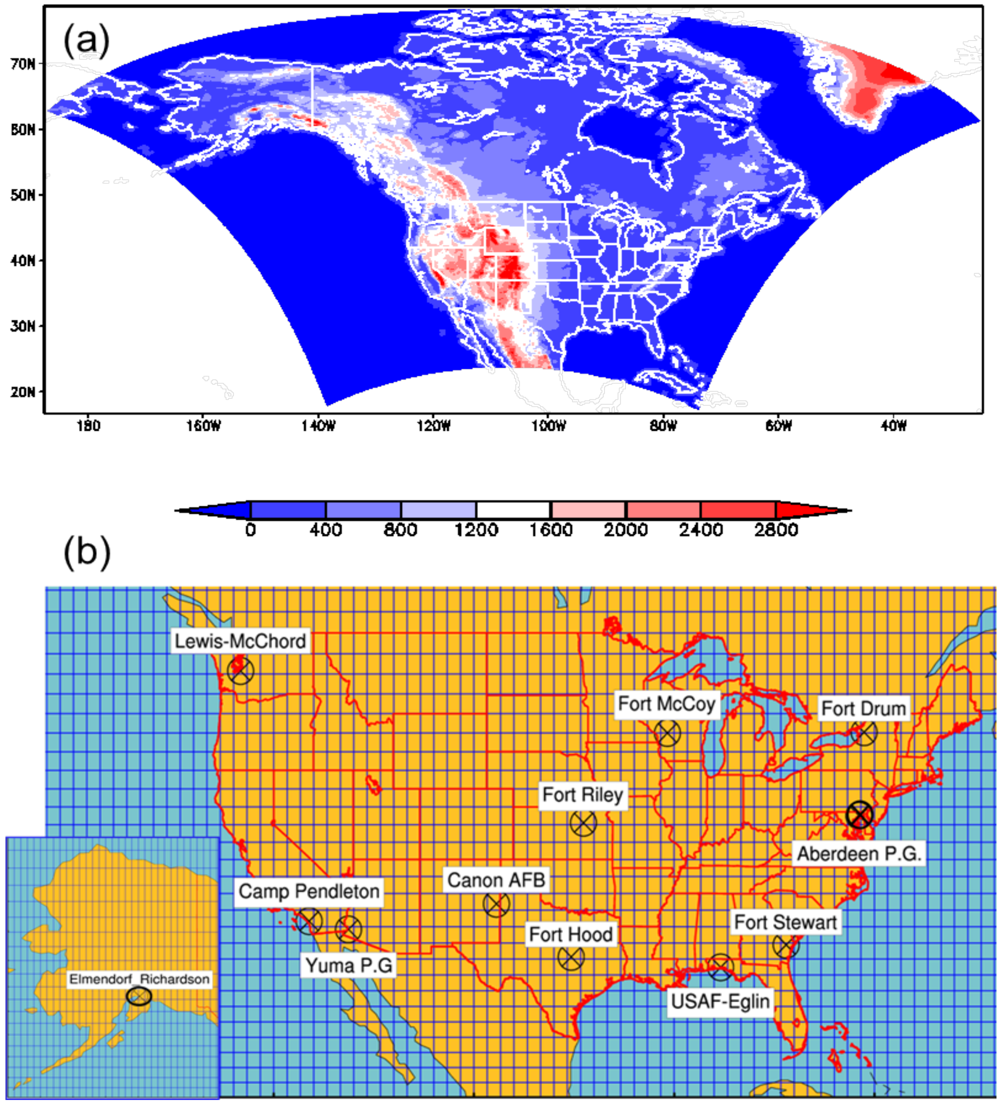

FIGURE 2 (a) Model Domain Used for the WRF Model Dynamic Downscaling Calculations, with Terrain Height Shaded; (b) DoD Installations Selected for Detailed Downscaling Model Performance and Evaluation

the accuracy of model projections or dealing with the uncertainty that comes with probabilistic information. Hindrances to use—such as scientific uncertainty, the lack of integrative models (hydrologic, fire, etc.) that use climate change estimates in risk analysis decision processes, and the lack of support from others at the installation—were noted by a few stakeholders.

Personnel involved with this aspect of the project would have benefited by meeting face to face with participants to discuss the questionnaire. The answers indicated that a number of stakeholders did not clearly understand the differences between weather forecasts and long-term climate change model projections. Higher-level decision makers should have been involved in these discussions. Future assessments of DoD stakeholders need to budget more time and

resources to enhance the exchange of knowledge between scientists and users. A detailed report from this activity is in Appendix A.

3.3 SELECTION OF REGIONS FOR MODEL EVALUATION

We considered several different ways of dividing the regions of the North America. The NARCCAP program proposed 31 regions covering the domain used in its simulations (Bukovsky 2011). These ecologically similar regions were identified to assist in installing NEON (National Ecological Observatory Network) flux sites and were built on a classification developed by Ricketts et al. (1999). Because data are generally insufficient to evaluate model performance in many of the 31 identified regions, we aggregated these regions into 10 larger regions (Figure 3) for our dynamic and statistical downscaling evaluations. For most of our analysis, the CONUS is broken into seven regions that are consistent with those used in the U.S. National Climate Assessment (Melillo et al. 2014). They are Northwest, Northern Great Plains, Southern Great Plains, Midwest, Northeast, Southwest, and Southeast (see Figure 2 in Janssen et al. 2013.

3.4 DATASETS FOR MODEL EVALUATION

We evaluated six datasets (Table 1) for calculating model bias for the dynamic and statistical downscaling techniques, including three surface air temperature datasets (Precipitation-Elevation Regressions on Independent Slopes Model [PRISM], Climatic Research Unit [CRU], and University of Delaware [UDEL]) and five precipitation datasets (PRISM, National Oceanic and Atmospheric Administration [NOAA] Climate Prediction Center [CPC], UDEL, CRU, North American Regional Reanalysis [NARR], and National Aeronautics and Space Administration [NASA] Tropical Rainfall Measuring Mission [TRMM]).

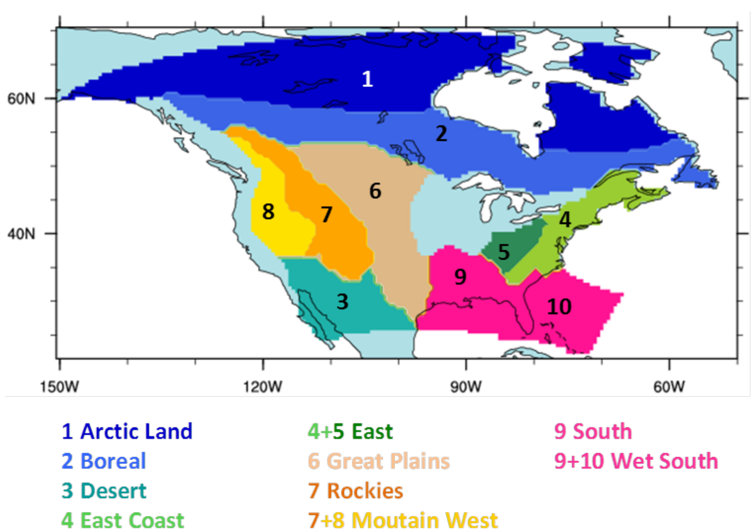

FIGURE 3 Regionalization Used in Our Model Evaluations

TABLE 1 Evaluation Datasets Applied in Our Studies

Name	Developer	Spatial and Temporal Resolution	Domain
PRISM	Oregon State University	4 km, monthly	CONUS
CPC	NOAA	0.25° × 0.25°, daily	CONUS
UDEL	University of Delaware	0.5° × 0.5°, monthly	Global
CRU	University of East Anglia (UK)	0.5° × 0.5°, monthly	Global
NARR	NCEP	32 km, 3 hourly	North America
TRMM	NASA	0.25° × 0.25°, 3 hourly	50°S–50°N; 180°W–180°E

The calculated bias in the model is different for each evaluation dataset (Figure 4). Through an extensive analysis of the observational datasets, we chose monthly temperature and precipitation from PRISM (developed by Daly et al. 1994, 1997, 2008) to evaluate the model's performance on the annual cycle. The PRISM values, which are corrected for systematic elevation effects on precipitation climatology, provide observation-based temperature and precipitation on a grid mesh of 1/8° latitude × 1/8° longitude that covers the entire CONUS. Given the strong dependence of temperature and precipitation on elevation, the topographic adjustment was critical, because cooperative stations over mountainous regions were preferentially located at lower elevations and thus tended to underestimate (overestimate) the true spatial average of precipitation (temperature). Therefore, the observation-based PRISM provides the most accurate data for precipitation and temperature, especially in topographically complex regions, and is the best available dataset for evaluating our high-resolution model simulations. Other widely used datasets such as UDEL and CRU are not necessarily able to resolve information that is as highly detailed as that from the WRF model used in this study; these other datasets are not discussed further.

Daily precipitation from the NOAA CPC at 0.25° × 0.25° (Chen et al. 2008; Xie et al. 2007) and daily temperature from the NCEP NARR (assimilated by observation; Mesinger et al. 2006; Bukovsky and Karoly 2007) were used to evaluate the model's daily performance in PDF. In addition, NARR 3-hourly precipitation was applied to evaluate the model's performance in the diurnal cycle of precipitation, which is the best available gridded dataset at the diurnal scale.

Based on these studies, we employ NARR data (Mesinger et al. 2006; Bukovsky and Karoly 2007) to evaluate model performance in near-surface relative humidity, wind, and high-level fields, such as geopotential height, humidity, and wind. The NARR assimilates observed information from multiple sources (aircraft, satellite, stations, etc.; see Tables 1 and 2 in Mesinger et al. 2006), and has been used widely as reference data by the climate downscaling community (e.g., Bowden et al. 2012; Otte et al. 2012; Liu et al. 2012; Loikith et al. 2013), although inaccuracies remain in some regions. For example, Bukovsky and Karoly (2007) found that, while the NARR provides a fairly good representation of observed precipitation over much of the CONUS, some inaccuracies appear over Canada because of the relatively poor data quality that NARR assimilates. Wang et al. (2016) found that NARR overestimates (underestimates) the

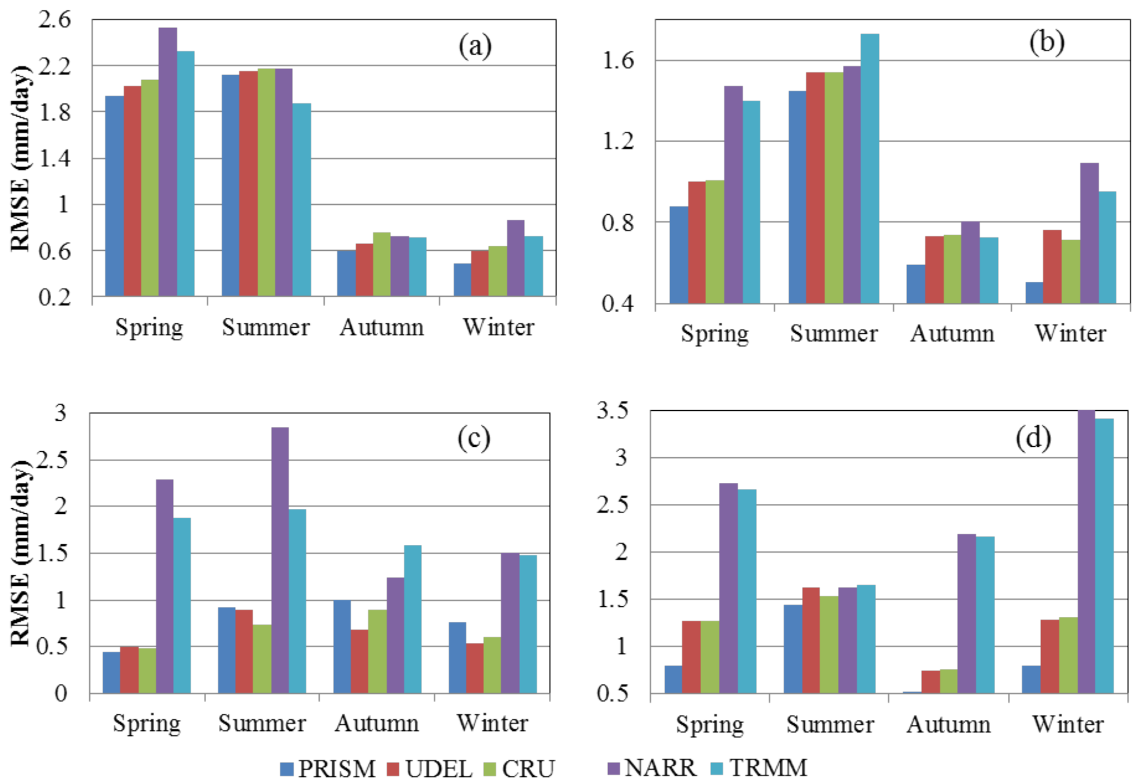

FIGURE 4 Regionally Averaged Seasonal Root-mean-square Errors (RMSEs) between the WRF Simulation and Five Validation Precipitation Datasets (PRISM, UDEL, CRU, NARR, and TRMM) over the (a) Great Plains, (b) Desert, (c) South, and (d) Rockies—Subregions 6, 3, 9, and 7, Respectively, in Figure 3

warming trend of January temperature over southeastern CONUS (over most of the western CONUS).

For other near-surface fields such as daily maximum and minimum temperature and precipitation, we use an observation-based gridded dataset that was constructed and documented well by Maurer et al. (2002). This gridded dataset has been applied extensively as meteorological references for evaluating dynamical and/or statistical downscaled results (e.g., Wood et al. 2004; Christensen et al. 2004; Maurer and Hidalgo 2008; Gutowski et al. 2010; Wehner 2013). The gridded precipitation within the CONUS is from the NOAA Cooperative Observer (Co-op) stations. The precipitation gauge data are first gridded to one-eighth-degree resolution using the synergraphic mapping system algorithm of Shepard (1984) as implemented by Widmann and Bretherton (2000). The gridded daily precipitation data are then scaled to match the long-term average of the parameter-elevation regressions on independent slopes model (PRISM) precipitation climatology (Daly et al. 1994, 1997), which is a comprehensive dataset that is statistically adjusted to capture local variations due to complex terrain. In this study, we also use the PRISM monthly precipitation dataset as the reference data to evaluate the model and understand the uncertainty of the model's performance relative to different reference data. The minimum and maximum daily temperature data over the CONUS, also obtained from Co-op

stations, are gridded using the same algorithm as for precipitation, and are lapsed to the grid cell mean elevation.

3.5 DYNAMICAL DOWNSCALING MODEL AND SIMULATIONS PERFORMED

The WRF model version 3.3.1 is applied at a horizontal resolution of 12 km, with 600 west–east and 515 south–north grid points over most of North America (Figure 2a). The lateral boundary conditions are specified in two different ways. As shown in Table 2, in the first set of the experiments, the WRF model is driven by the reanalysis of the NCEP-R2 (National Centers for Environmental Prediction—U.S. Department of Energy Reanalysis II) over the period 1980–2010. In the second through sixth sets of experiments, the WRF models are driven by datasets from three fully coupled GCMs. These five sets of experiments span three different time periods: 11 years over the historical period (1994–2004), 11 years over the mid-21st century (2044–2054), and 11 years over the late 21st century (2084–2094). The forcing scenarios for these future simulations in both the global and regional climate models are from CMIP5. The name of each GCM dataset is listed in Table 2. More details of the model are presented in Wang and Kotamarthi (2015).

The six WRF model runs listed in Table 2 have the same horizontal resolution. They are also the same in most of their "physics," which include the Grell-Devenyi convective parameterization (Grell and Devenyi 2002), the Yonsei University planetary boundary layer

TABLE 2 Dynamical Downscaling Simulations Completed by Argonne and University of Illinois at Urbana-Champaign, Funded by SERDP

Boundary Conditions[a]	Spectral Nudging	Nudging Strength	Periods	Spin-up Time	Scenarios
NCEP-R2	Yes	$3 \times 10^{-4}\ s^{-1}$	1980–2010	1 day	NA
CCSM4 (raw)	Yes	$3 \times 10^{-5}\ s^{-1}$	1995–2004 2085–2094	1 year	RCP 4.5/8.5
CCSM4 (bias corrected)	Yes	$3 \times 0^{-5}\ s^{-1}$	1995–2004 2045–2054 2085–2094	1 year	RCP 4.5/8.5
GFDL-ESM2G (bias corrected)	No	NA	1995–2004 2085–2094	1 year	RCP 8.5
GFDL-ESM2G (bias corrected)	Yes	$3 \times 10^{-5}\ s^{-1}$	1995–2004 2085–2094	1 year	RCP 8.5
HadGEM-ES (raw)	No	NA	1995–2004 2085–2094	1 year	RCP 8.5

[a] CCSM4 = Community Climate System Model, version 4; GFDL ESM2G = Geophysical Fluid Dynamics Laboratory Earth System Model with Generalized Ocean Layer Dynamics component; HadGEM-ES = Hadley Centre Global Environment Model, version 2-Earth System; NCEP-R2 = National Centers for Environmental Prediction—U.S. Department of Energy Reanalysis II.

scheme (Noh et al. 2003), the Noah land surface model (Chen and Dudhia 2001), and the longwave and shortwave radiative schemes of the Rapid Radiation Transfer Model for GCM applications (see http://rtweb.aer.com; Iacono et al. 2008). However, as shown in Table 2, the first WRF run, which was driven by NCEP-R2, uses WSM6 (WRF Single-Moment 6-Class) microphysics and applies spectral nudging with a nudging coefficient of 3×10^{-4} s^{-1}. Moreover, it only allows 1 day for spin-up time and is re-initialized every year. Because the NCEP-R2 driven run was conducted first, we found some reasons for the model bias were nudging strength and choice of microphysics (Section 3.5; also see Wang and Kotamarthi 2013). Therefore, we adjust these settings for the GCM-driven runs. As discussed in Section 4.4.1, we also conducted sensitivity experiments and found that using weaker nudging, Morrison microphysics, and longer spin-up time helps reduce several different aspects of the model bias. Thus, for some of the WRF runs, which are driven by various GCMs, we use the Morrison microphysics scheme (Morrison et al. 2009) and 1-year spin-up time for each 10-year continuous run. For those runs that apply spectral nudging, the nudging coefficient is 3×10^{-5} s^{-1}. When applying spectral nudging, we are nudging selected model calculated fields to those derived from the GCM boundary conditions, but bias corrections are applied to the boundary conditions based on the climatology and not the 6-hourly variations.

This study compares WRF runs driven by the Community Climate System Model, version 4 (CCSM4) developed by National Center for Atmospheric Research, United States (Gent et al. 2011), WRF runs driven by the Geophysical Fluid Dynamics Laboratory Earth System Model with Generalized Ocean Layer Dynamics component (GFDL-ESG2G) developed by the NOAA/Geophysical Fluid Dynamics Laboratory, United States (Donner et al. 2011), and the Hadley Centre Global Environment Model, version 2-Earth System (HadGEM2-ES) developed by the Met Office Hadley Centre, United Kingdom (Jones et al. 2011). These three GCMs show different climate sensitivities when forced by doubled atmospheric carbon dioxide concentration, with CCSM4 showing 2.92 K of warming effect, GFDL-ESM2G showing 2.38 K of warming effect, and HadGEM2-ES showing 4.55 K of warming effect (Sherwood et al. 2014). To explore the impacts of spectral nudging on model performance when bias correction is applied, we conducted two WRF runs driven by GFDL-ESG2G, with spectral nudging turned on in one of the simulations and turned off in the other simulation. We created a database of model simulations at a spatial resolution of 12 km using these GCMs as boundary conditions. The simulations we have performed are summarized in Table 2 and available with Argonne National Laboratory. The temporal resolution of the model output is 3 hours for all the simulations, and includes more than 50 variables produced by the model, resulting in a data volume of nearly 200 Tb. These data are stored as a series of NetCDF files for each day of the simulation.

3.6 STATISTICAL DOWNSCALING MODEL

For this project, we use the Asynchronous Regional Regression Model (ARRM; Stoner et al. 2012). This model was selected for this second task because it can resolve the tails of the distribution of daily temperature and precipitation to a greater extent than the more commonly used Delta and Bias Correction-Statistical Downscaling methods (e.g., Hay et al. 2000; Wood et al. 2004). This method is far less time intensive compared to an RCM such as WRF, and therefore is more accessible as well as less costly; however, it cannot produce the large suite of

output variables an RCM can compute, only the more frequently studied variables such as precipitation, minimum and maximum temperature, solar radiation, and relative humidity.

ARRM is based on the statistical concept of asynchronous quantile regression, which was originally developed by Koenker and Bassett (1978) to estimate conditional quantiles of the response variable in econometrics. This regression is asynchronous; data values that are regressed against each other did not necessarily occur the same calendar day, but rather correspond by quantile or rank. The regression model derived from historical model simulations and observations can then be applied to future model simulations, to project downscaled future conditions.

Dettinger et al. (2004) were first to apply this statistical technique to climate projections to examine simulated hydrologic responses to climate variations and change, as well as to heat-related impacts on health (Hayhoe et al. 2004); subsequent versions of this algorithm have been used in city-scale projections for the Northeast Climate Impacts Assessment (Frumhoff et al. 2007) and the Chicago Climate Action Plan (Hayhoe et al. 2010b), as well as in the upcoming 2014 Third U.S. National Climate Assessment (Walsh et al. 2013). The version used by Dettinger et al. (2004) assumes a linear relationship between the modeled and observed quantiles, assuming the variable being downscaled has a normal distribution. However, the ARRM model allows for nonlinear relationships by applying a piecewise linear regression function to each of the 12 months of the year, which allows it to capture more of the variation in the distribution of the variable of interest.

4 RESULTS AND DISCUSSIONS

We have performed a comprehensive analysis of the model output generated by this project and other climate model output available in various data archives developed by the climate modeling community. We describe the methodology used for performing these analysis, metrics used for evaluating the model performance and the suggested use of the model data archive we have generated.

4.1 DYNAMICAL DOWNSCALING SIMULATIONS AND BIASES

We evaluated of the model bias extensively through 30 years of simulations (1980 to 2010) with WRF in a climate simulation mode. We use a regional-scale model at a spatial resolution of 12 km covering most of North America (Figure 1) for regional climate simulations. The domain has approximately 14 million grid cells (600 × 516 × 38). Output is saved every 3 hours, which adds up to about 6.8 GB/day (0.86 GB/hour). The total simulation of 70 years included 30 historical years and 40 future years (time slices for RCP4.5 and RCP8.5 in 2045–2054 and 2085–2094). The simulations were performed on the Argonne Leadership Computing Facility (ALCF) flagship computing cluster (Intrepid/Mira) and on the National Energy Research Scientific Computing Center (NERSC) cluster (Hopper). An optimal model configuration was established through sensitivity studies (Wang and Kotamarthi 2014).

Because the primary evaluation dataset (PRISM) was for CONUS regions, the model evaluation region used in this study was also confined to CONUS. Four subregions—Great

Plains, South, Desert, and Rockies (Figure 3)—with various model biases and climatic and topographic features were selected for our main analysis. To retain the original features of the datasets at the resolutions during model evaluation and for comparison with other datasets, we used the native resolution (with no interpolation) of the datasets/models to plot geographic patterns and calculated statistical metrics except for spatial and temporal correlation coefficients, which were computed by re-gridding the PRISM data from 4 km to the WRF at 12 km and to the NARCCAP WRFG (Weather Research and Forecasting Grell) at 50 km.

4.1.1 Surface Air Temperature

The WRF model's ability to capture the annual cycle in air temperature is assessed by comparing it with PRISM data and the NARCCAP-WRFG simulation—one of the best available dynamic downscaled products widely used by the community. In addition, the WRF model's ability to retain and add values above its driver is evaluated by comparing it with National Centers for Environmental Prediction-U.S. Department of Energy Reanalysis II (NCEP-R2) data. Figures 5a–5d compare the subregional average bias of temperature for the WRF, NARCCAP-WRFG, and NCEP-R2 versus PRISM in all four seasons from 1980 to 2004. The error bars denote the yearly distribution of the biases at the 10th and 90th percentiles. Generally, the annual variations in the biases for WRF and NCEP-R2 are smaller than those for NARCCAP-WRFG. Improvements of WRF over NARCCAP-WRFG are seen over the Great Plains (Figure 5a) in spring and winter, over the Desert (Figure 5b) in all four seasons, and over the Rockies (Figure 5d) in spring, summer, and winter. For example, over Desert, the NARCCAP-WRFG bias was 2.8–4.4°C, while the WRF bias was only 0.06–1.0°C; over the Rockies, the NARCCAP-WRFG bias was 0.3–2.6°C, while the WRF bias was 0.2–2.0°C (except in fall). Topography plays a key role in determining temperature over mountain ranges, and (as expected) the WRF model with its better representation of topography leads to smaller temperature RMSEs than the NARCCAP-WRFG over the Desert and Rockies. For example, over the Desert region, the RMSEs for NARCCAP-WRFG were 2.9–4.5°C, while those for WRF were only 0.3–1.2°C; over the Rockies, the RMSEs for NARCCAP-WRFG were 0.8–2.8°C, while those for WRF were 0.4–2.1°C. However, WRF shows larger warm biases than NARCCAP-WRFG over the Great Plains (Figure 5a) in summer and fall and over Southern Central (SC) regions (Figure 5c) in all four seasons. The RMSEs for WRF over SC were also larger than those for NARCCAP-WRFG in all seasons except summer.

To achieve good fidelity for the WRF model in simulating climate, we expect the model not only to capture the mean fields and reduce model bias, but also to generate variances similar to those of the observations. Thus, in addition to assessing mean fields, we compared the standard deviation (SD) of temperature for WRF with the SD of the PRISM observations, as

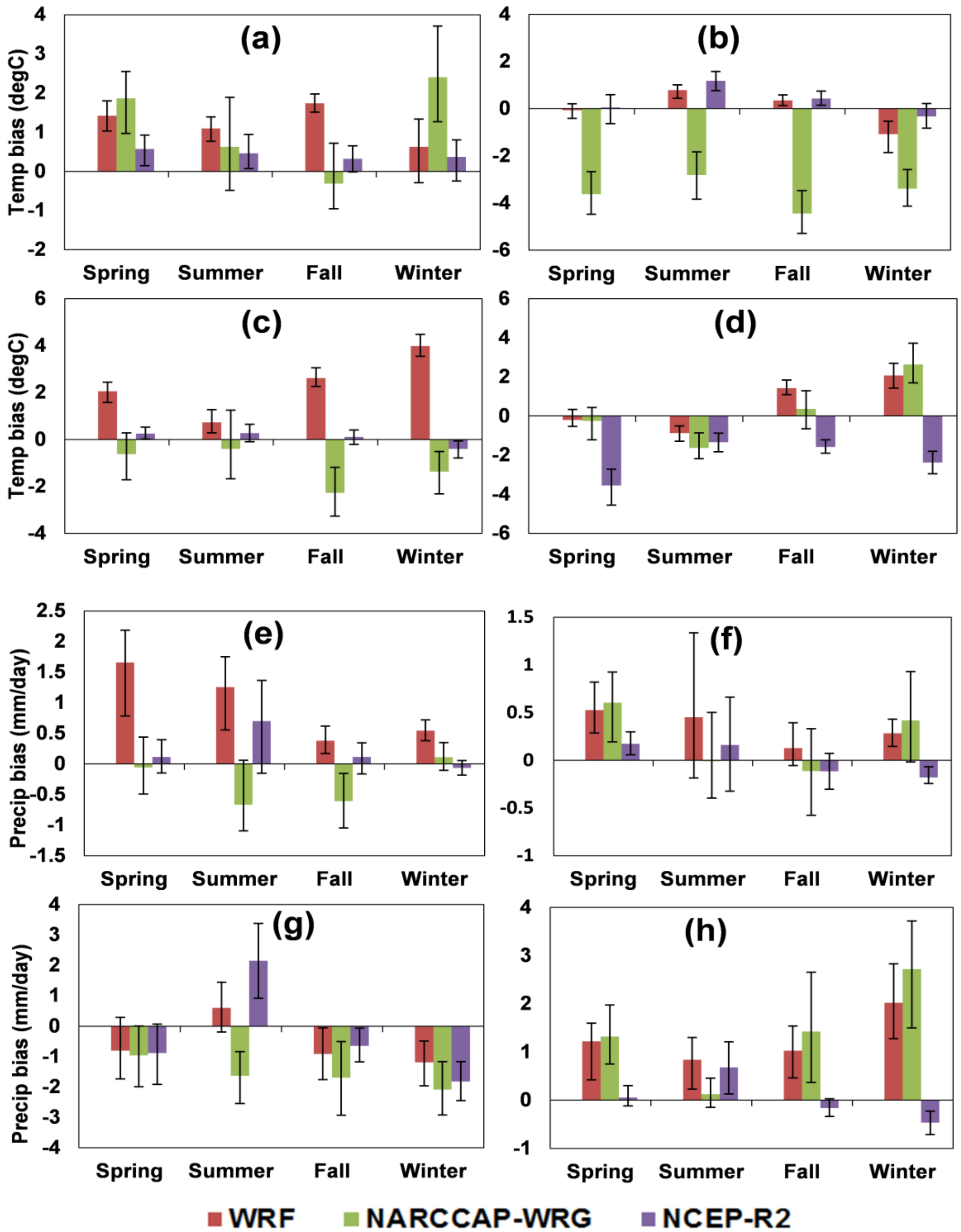

FIGURE 5 Subregional-average Bias of 2-m Temperature (a–d) and Precipitation (e–h) from WRF, NARCCAP-WRFG, and NCEP-R2 versus PRISM Data by Season in 1980–2004 over the Great Plains (a, e), Desert (b, f), South (c, g), and Rockies (d, h); Error Bars Denote the Annual Distribution of Bias at the 10th and 90th Percentiles

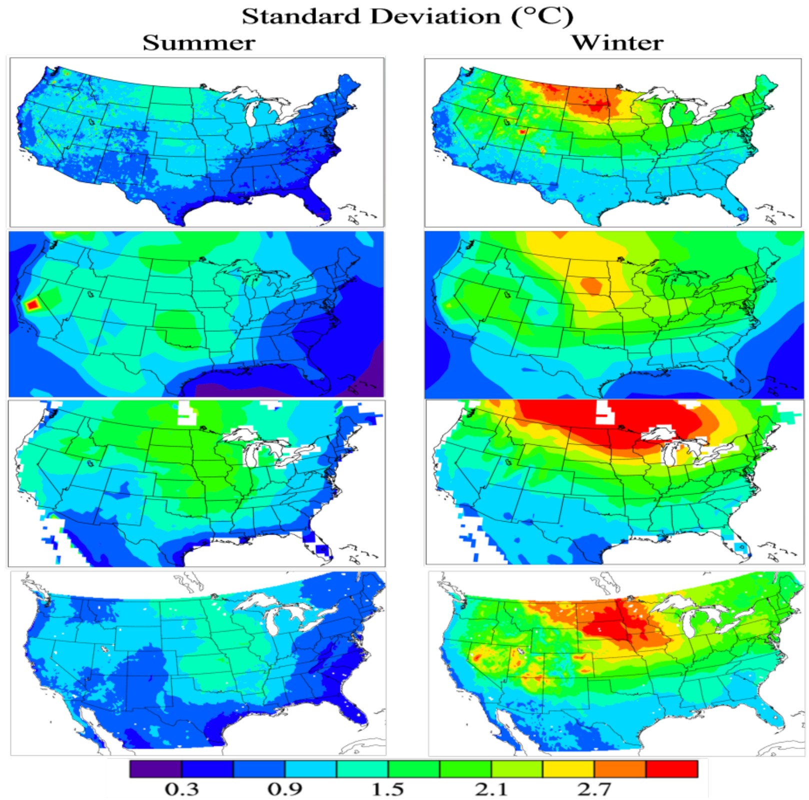

FIGURE 6 Multi-annual (1980–2004) Average Summer and Winter Temperature Standard Deviation from PRISM (top row), NCEP-R2 (second row), NARCCAP-WRFG (third row), and WRF (bottom row)

shown in Figure 6. The WRF captured the spatial pattern and the value of the temperature SD very well in each season, with the smallest SD in summer and the largest SD in winter.

4.1.2 Precipitation

Figure 7 compares the monthly variations of precipitation between PRISM, NCEP-R2, NARCCAP-WRFG, and WRF over four subregions: Great Plains (Figure 7a), Desert (Figure 7b), South (Figure 7c), and Rockies (Figure 7d). Like the PRISM observations, WRF generated a rainfall peak in May and June over the Great Plains and in July and August over the Desert (Figure 7b). This similarity in the monthly transition in precipitation is encouraging, because it is an important indicator of the pre-North American Monsoon (NAM) and the onset and peak of NAM (Castro et al. 2007). However, over the Great Plains (Figure 7a), WRF produced much heavier rainfall than the PRISM values from January to July. Over the South (Figure 7c), the NCEP-R2 generated substantial wet (dry) bias in the summer (winter) months—opposite the PRISM values. NARCCAP-WRFG presents a monthly variation similar to that of

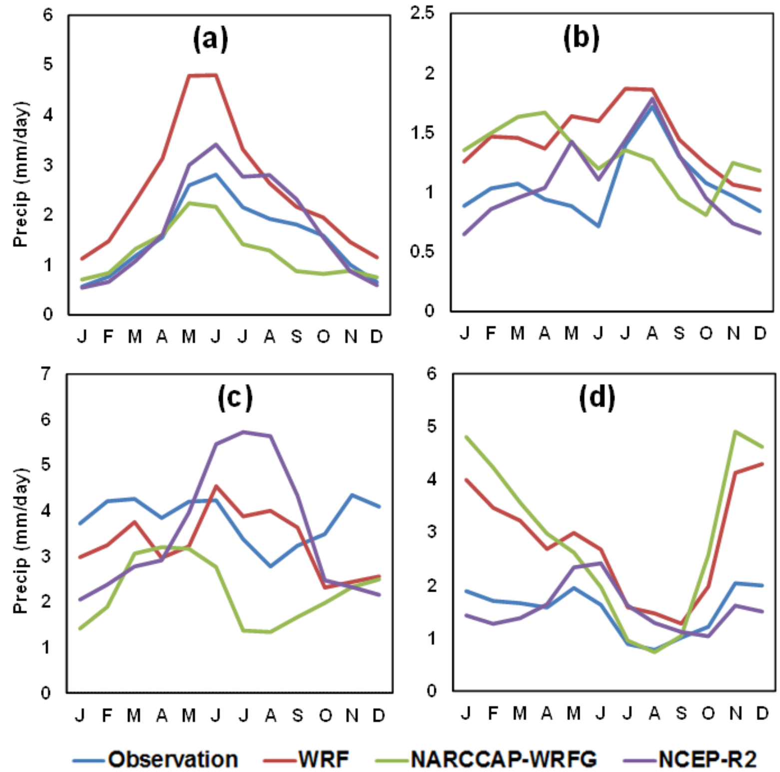

FIGURE 7 Multi-annual (1980–2004) Average Monthly Variations in Precipitation over (a) Great Plains, (b) Desert, (c) South, and (d) Rockies from PRISM, WRF, NARCCAP-WRFG, and NCEP-R2

PRISM but shows a significant dry bias during the entire year. Especially in the period from January to March, WRF produced heavier rainfall than NARCCAP-WRFG, closer to the PRISM values. However, WRF generated less (more) rainfall in the cold (warm) season than the PRISM values, possibly because of the lateral boundary conditions (LBCs) of NCEP-R2. Over the Rockies (Figure 7d), both NARCCAP-WRFG and WRF show wet biases in the cold season, when the region receives most of its annual precipitation. WRF shows slightly better monthly variation than NARCCAP-WRFG during the cold season, with less rainfall (closer to the PRISM values).

The temporal correlation coefficients (TCCs) between simulated and observed monthly precipitation over each grid were significantly improved by WRF versus NARCCAP-WRFG (Figure 8). WRF captured the monthly variations in precipitation over the CONUS, with TCC > 0.5 (significant level = 0.005) over most of the CONUS, except the Rockies. NARCCAP-WRFG shows significant TCC values over the western CONUS, but the lower TCC values over the eastern CONUS, especially close to the ocean and the Gulf of Mexico, do not pass the statistical significance test.

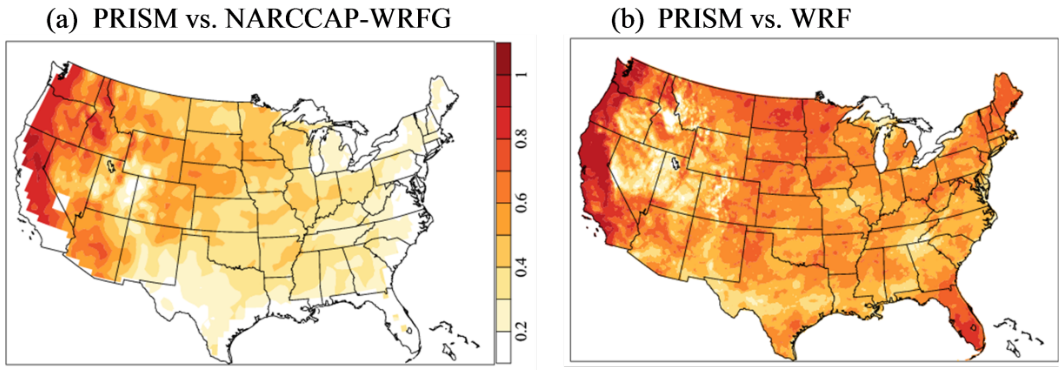

FIGURE 8 Temporal Correlation Coefficients in Precipitation (a) between PRISM and NARCCAP-WRFG at 50 km and (b) between PRISM and WRF at 12 km during all months in 1980–2004 (The student t-test result at the 0.005 [and lower] level of significance is marked by the color scale.)

4.2 A GENERAL EVALUATION OF PERFORMANCE FOR THE ENTIRE ENSEMBLE

We apply Taylor diagrams (Taylor 2001) to evaluate the models' performance in daily near-surface variables. A Taylor diagram concisely relates a model and a reference dataset pattern correlation, their root-mean-square error (RMSE), and standard deviations (STDEV) (Taylor 2001). Figures 9a–9c display a set of Taylor diagrams that show several near-surface variables: daily maximum temperature, minimum temperature, and precipitation, using the script provided by NCAR Command Languages to visualize these results (https://www.ncl.ucar.edu/Applications/taylor.shtml). STDEV is calculated as a ratio of model STDEV over the reference dataset's STDEV, so the closer to 1 ("REF") the values are, the better the model is at capturing the spatial and temporal variance. The correlation coefficients are plotted along the semicircle along the outermost part of the graph. We calculate correlation and variance across not only latitude and longitude, but also the time dimension. In general, the WRF simulation driven by NCEP-R2 (open circles) performs significantly better in pattern correlation than the WRF simulations driven by GCMs for all regions and predicts the STDEV with overall more accuracy than the mean (crosses) and median runs (asterisks). The mean and median have slightly higher correlations than the GCM driven simulations for maximum and minimum temperatures, but they fall to the left of "REF" line in most regions, which means they underrepresent the temporal and spatial variations of temperature. Although it is hard to categorically rank those simulations from a Taylor diagram driven by different GCMs, it is clear that the accuracy of a model simulation depends on the fields and the regions of interest. For maximum and minimum temperature, as shown in Figures 9a–9b, in each subregion, most of the WRF simulations driven by different GCMs are grouped together; some yield higher correlation or STDEV values than others, but how those models rank regionally in terms of error will be discussed in Section 3.2. For precipitation (Figure 9c), all the WRF simulations (including mean and median) show much less skill in predicting pattern correlation than for the temperature variables. The STDEV values in the daily precipitation is variable across the GCM driven ensemble for all the regions, but their STDEV ratio remains between 1.25 and 0.75. The mean and median have a much greater underestimation of STDEV for precipitation with values less than half of the observed STDEV.

4.2.1 Relative Error

Taylor diagrams consider the temporal and spatial pattern recognition, but they do not define a model's bias. They also fail to decipher the model's ability to observe the tails of the observed PDF, which is one of the important advantages the downscaled GCM runs have over the raw GCM output. Here and in Section 3.7.1, we evaluate model performance based on metrics that describes relative error of daily mean and PDF that drawing distribution tails.

We employ the performance metrics developed by Gleckler et al. (2008). To begin, RMSE is calculated for each variable of interest and NCA subregion for all six model runs, as well as the mean and median in the ensemble. The reference dataset depends on the type of variable being analyzed. NARR is used to evaluate above surface variables (e.g., Liu et al. 2011) while the gridded observations are used to evaluate the appropriate surface variables (e.g., temperature and precipitation). As shown in equation (1), to calculate relative error for a field f (E_{mf}'), we define a typical model error ($\overline{E_f}$), which is the median of RMSEs for the six simulations plus the median and mean. We use the median of RMSEs rather than the mean as the typical model error to prevent models with unusually large errors from disproportionately influencing the results (Gleckler et al. 2008). $\overline{E_{mf}}$ is the RMSE of one particular simulation out of six simulations, plus the mean or median. The relative error is a measure of how well the particular model of interest performs compared to the typical model error in the ensemble. For example, if a model has a negative $\overline{E_{mf}}$, this means it has smaller RMSEs than the simulations with positive $\overline{E_{mf}}$:

$$E'_{mf} = \frac{E_{mf} - \overline{E_f}}{E_f} \tag{1}$$

Figure 10 shows the relative error for daily precipitation (Figure 10a), mean temperature (Figure 10b), and daily maximum/minimum temperature (Figures 10c and 10d) over seven NCA regions and CONUS from the WRF simulations comparing with the gridded observation dataset described in earlier. In general, the WRF simulations driven by GCMs score worse for all four variables than the ensemble mean and median. For precipitation, the WRF_HadGEM and WRF_GFDLNN show less RMSE than other WRF simulations driven by GCMs in the Midwest and that includes the NCEP-driven simulation for the Northern Great Plains and Southern Great Plains regions. The WRF_NCEP and WRF_CCSM_nBC predict lower RMSEs than other WRF simulations driven by GCMs in the Northeast and Southeast regions. There are noticeable differences between the models with and without bias correction. The relative error between WRF_CCSM_nBC and WRF_CCSM_BC has the greatest differences for precipitation in the Northern Great Plains, Southern Great Plains, and Midwest. Using bias correction for these regions caused larger bias than when no bias correction is applied to the boundary conditions. A similar trend is observed for models with and without spectral nudging. For example, WRF_GFDLN shows larger bias in precipitation relative error than does WRF_GFDLNN run for all regions except for the Northeast.

It is worth mentioning that over the Great Plains, the WRF_NCEP shows positive relative errors, which means it has larger RMSEs than the typical model error, and even larger RMSEs

than some of the GCM-driven runs. This is because, although we are using the “perfect” boundary conditions, the physics and the model setup are somewhat different from the other WRF simulations driven by GCMs. The WRF_NCEP run was the first run we conducted for the project, aiming to understand the model performance and the model sensitivity to different physics and setup. In this run, we only allowed 1 day for spin-up time, and we re-initialized the model every year. These are two of the reasons that model shows a wet bias over the Great Plains. In addition, the microphysics scheme that was applied for the run also induces a wet bias over the Great Plains in cold seasons. Thus, we modify the model setup and microphysics for WRF simulations driven by GCMs to reduce the bias generated by those factors.

The preferred GCM and the needed corrections made to the boundary conditions for mean temperature (Figure 10b) and maximum temperature (Figure 10c) is regionally and simulation dependent. For minimum temperature (Figure 10d), the WRF_GFDLNN shows smaller relative errors than does WRF_GFDLN for all the regions. WRF_HadGEM shows the lowest relative error in comparison with other WRF simulations for all but two regions: the Midwest and the Northeast. There is not much difference between WRF_CCSM_BC and WRF_CCSM_nBC, but spectral nudging in the WRF_GFDLN run leads to much higher RMSEs than the lack of spectral nudging shown in the WRF_GFDLNN runs. It is worth noting that the GCM-driven simulation that has the lowest error for every region, other than the Midwest, is one that does not employ spectral nudging. The CCSM4 runs both used spectral nudging and showed far less error than the WRF_GFDLN run. For minimum temperatures, WRF_CCSM_BC was significantly more accurate for all eight regions than the WRF_GFDLN. Since both simulations employ bias correction and nudging, much of the error in the GFDL minimum-temperature runs is likely due to the biases in the boundary conditions of that GCM. It would provide significant value if one could develop a single index to evaluate individual model performance considering all the variables of interest (Gleckler et al. 2008). This provides us a process that we could use to put more weight on the “better” model than the “worse” model when considering future climate projection. However, different model outputs are related to different aspects of model physics and/or model setup. Subjectively ranking the model performance based on an average score of all the variables of interest would substantially hide model errors for some aspects (Gleckler et al. 2008). In this study, we rank the model over each region for each variable. Users are allowed to give appropriate weights to different models depending on their real applications.

Figure 11 shows how using PRISMM (left) or NARR (right) as the reference datasets for monthly precipitation can produce different RMSEs. Overall, many of the regional ranks of the models are similar between the two, but there some of the ranks of the ensemble display slight differences. For example, the WRF_HadGEM performs the best in the southwest using NARR, but is second to WRF_NCEP when using PRISM. Over a domain as large as the CONUS (Fig 11, far left column) the differences in error between the reference data are not significant, but this study focuses most of its analysis on regional biases. There are several cases where the difference in the magnitude of the RMSE is as high as 25% (e.g., Northeast region for GFDL using nudging and Southwest region for HadGEM2). This error or uncertainty is hard to adjust or fix, but it still cannot be ignored. Similar to the GCM-driven runs, the observed monthly rainfall in the reference data used for this study could be accurate over long term, but could still fail to capture daily variations, particularly in mountainous areas (Alexander et al. 2006; Hijmans et al.

2005). Therefore, using multiple reference datasets that yield multiple results for errors for both relative

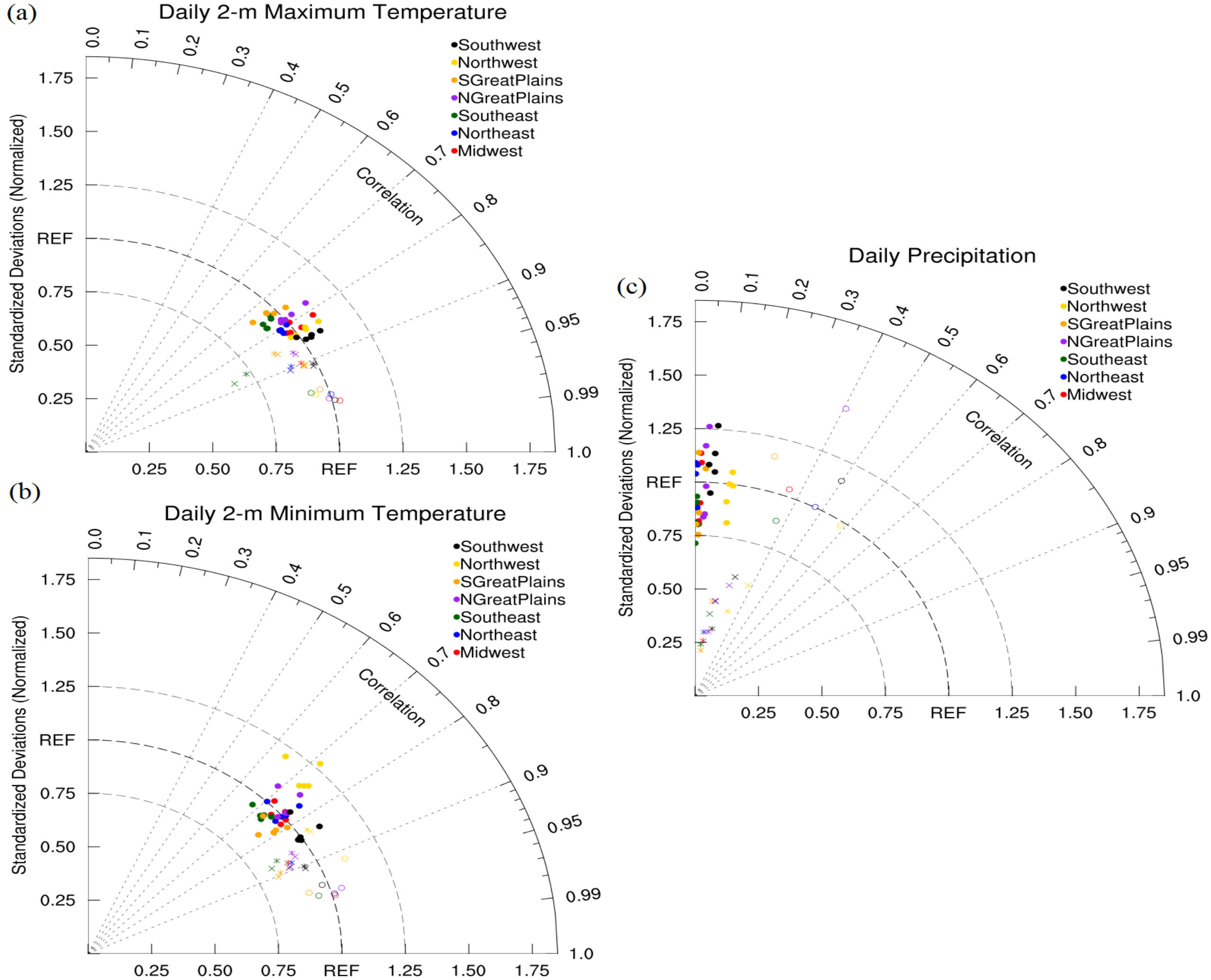

FIGURE 9 (a) Taylor Diagrams for Maximum Temperature, (b) Minimum Temperature, and (c) Precipitation (The filled circles represent the WRF simulations driven by different GCMs and the open circles are the NCEP-R2-driven WRF run. The crosses indicate the mean for the ensemble at each location and the asterisks are the medians.)

error and extremes for a historical period, provides a more comprehensive understanding of the model performance. Understanding where the simulations fail, or do not closely match the observations, is the most important feature of this research and is vital to understanding future projections of climate extremes (Ekström et al. 2005). Similarly, the WRF model itself and the physical schemes it uses introduce an additional set of regional biases (e.g., Ruiz et al. 2010; Jankov et al. 2005; Ries and Schlünzen 2009; Cheng and Steenburgh 2005; Aligo et al. 2007). All of these studies discuss the importance of WRF configuration and conclude that the ideal settings will have high temporal and geographical dependence based on the test variables. The large domain of this study evaluates regions with varying topographies and climates.

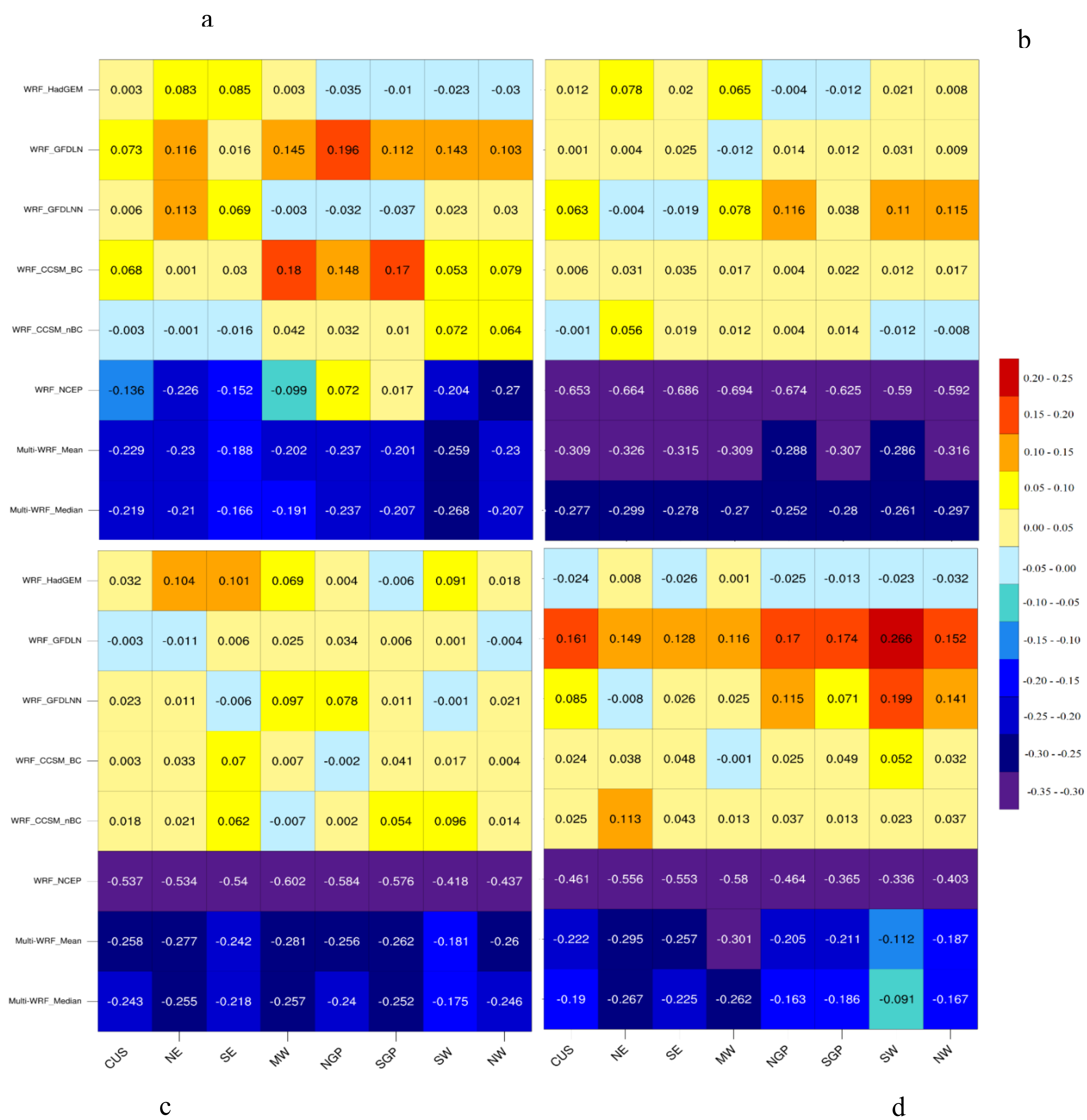

FIGURE 10 RMS Error for Surface Variables Compared to Observed Gridded Values (Top left: daily precipitation. Top right: daily mean temperature. Bottom left: daily maximum temperature. Bottom right: daily minimum temperature. The y-axis shows members of a model ensemble generated for the project and the x-axis shows different regions of the country [CUS—CONUS; NE—Northeast; SE—Southeast; MW—Midwest; NGP—Northern Great Plains; SGP—Southern Great Plains; SW—Southwest; NW—Northwest].)

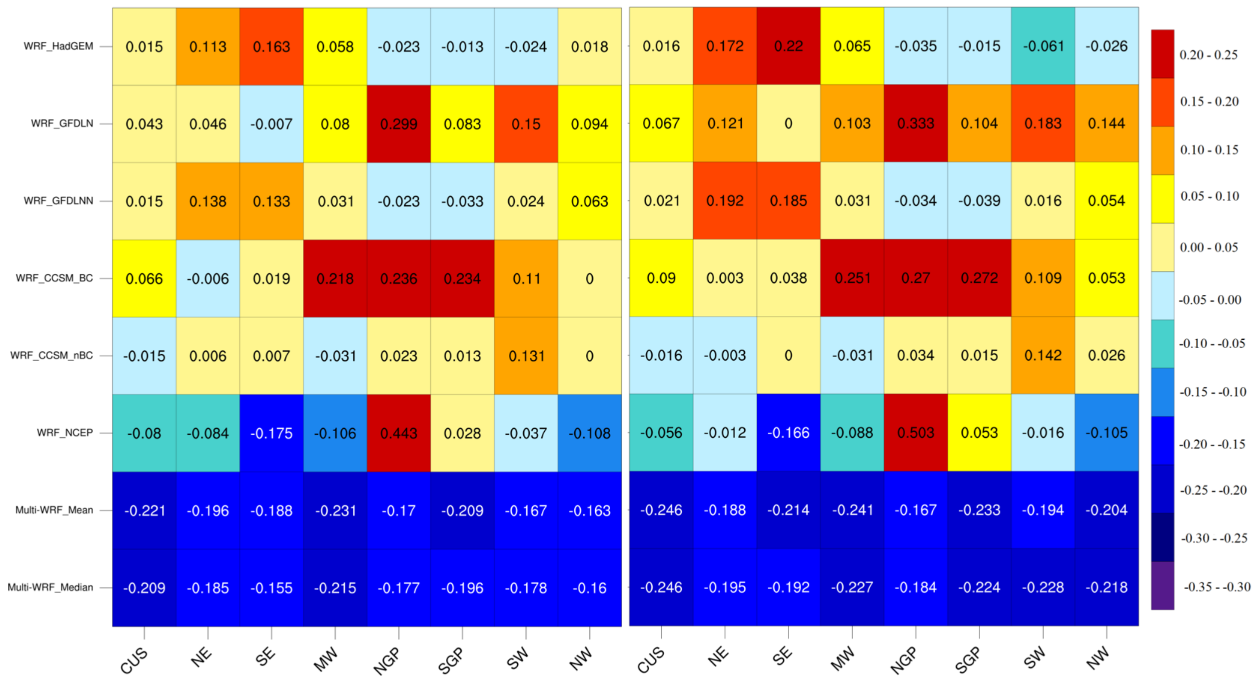

FIGURE 11 RMS Error for Total Monthly Rainfall Compared to Two Reference Datasets: PRISM (left) and NARR (right)

4.3 EXTREME EVENTS

Our survey (Section 3.2) identified a concern about extreme events at DoD installations and possible trends in their frequency and intensity. The primary motivation for dynamically downscaling climate models is to gain a more comprehensive idea of regional extreme events, and eventually make more accurate predictions of frequency and intensity of future anomalous climate events. In this section, we discuss the models' reliability at forecasting the frequency and/or intensity of extreme warm/cold temperatures, heat stress, single- and multi-day heavy precipitation events, and dry spells.

4.3.1 Temperature Extremes

Temperature values that are located in the right tail of the PDF curve provide valuable information about how the model simulation captures the extreme maximum temperature for a given location. In this study, we calculate the 95% threshold of the summer (June, July, and August) maximum near-surface temperature and the 5% threshold of the winter (December, January, and February) minimum near-surface temperature in the reference data and WRF simulations to judge the model's ability to capture these extreme high and low temperatures. Figure 12 shows the differences in extreme high temperature between model simulations and the observations based on NCA subregion averages. To calculate these values, first the 95% thresholds for each grid point in both observations and the simulations are calculated; second, the

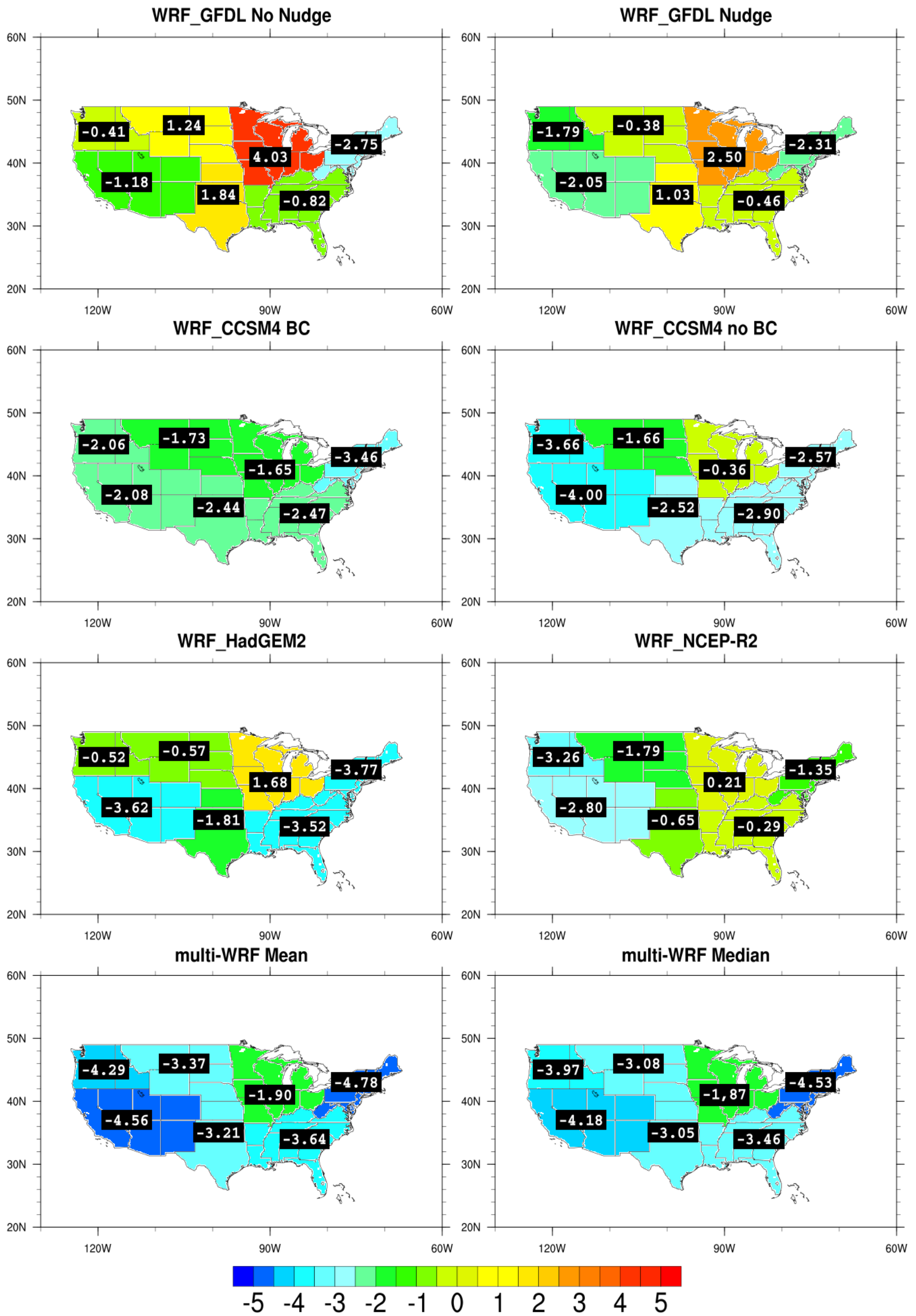

FIGURE 12 Average Regional Difference in 95% Threshold of Daily Maximum Summer (June, July, and August) Temperature (in °C) between the Models and Observations

differences are found for each grid point; and last, regional average of the differences are displayed. Compared with the six individual model output, the mean and median have the largest differences in extreme high temperature (cold bias) from the reference dataset in most of the regions because mean and median filters out the day-to-day variability at each location, which reduces the variance of the PDF curve and acts to smooth out the real extremes. Both CCSM4 WRF simulations (with and without bias correction) underestimate the maximum temperature extremes for all seven climate regions. The use of bias correction reduces the cold bias over the Northwest, Southwest, and Southeast regions by 0.5–2°C. However, using bias correction on the CCSM4 models increases the bias over the Midwest and Northeast regions by 1°C compared to the run without bias correction. Overall, the GFDL-driven simulations have warm bias over the Great Plains and Midwest and a smaller cold bias than the CCSM4 driven runs. In the two runs where WRF_GFDLN and WRF_CCSM4_BC use both spectral nudging and bias correction, there are large differences in all seven regions, indicating that the GCM used to force the WRF makes a larger difference than the use of bias correction and nudging does for maximum temperature. This is especially true for the Midwest and the Southern Great Plains, where the two runs not only disagree on sign, but the 95% thresholds differ by more than 3°C. Spectral nudging does improve the model performance in extreme high temperature over most of the regions. For example, the bias in the WRF_GFDLN run for the Northern Great Plains, Southern Great Plains, Midwest, Southeast, and Northeast regions is smaller by 0.36–1.53°C than in the WRF_GFDLNN.

Figure 13 shows the differences of 5% threshold of winter (December, January, and February) minimum temperature. Like maximum temperature, the tail in the PDF curve for GCM driven simulations are too close to the mean and underestimate the intensity of extreme cold temperature extremes for many of the regions. The only region where all the models (except WRF_NCEP) are consistently too cold is the Northwest region. The HadGEM2 model was the closest to the reference data set by almost 2 °C for this region when compared to the other GCM runs. Nudging and bias correction effect the GFDL and CCSM4 runs differently. Spectral nudging shifts the threshold value in WRF_GFDLN to the right for all the regions, making the extreme cold temperature closer to the observation than WRF_GFDLNN over Northwest, but further from the observation than WRF_GFDLNN over other regions. Different from the effects of nudging, bias correction reduces the bias of extreme cold temperature by as much 2 °C in most of the NCA subregions, with the exception of NW and SW. The two simulations (WRF_GFDLN and WRF_CCSM_BC) that use both of bias correction and nudging have the same sign and similar 5% threshold magnitudes with the exception of the magnitude for the Northeast. This is different than the maximum temperature differences, where the differences between the GCM boundary conditions were a much more significant factor in the biases between these two runs.

4.3.2 Heat Index

In addition to temperature, relative humidity plays an equally important role on the amount of stress the human body can endure in hot conditions. Thus, world-wide heat indices were developed (Buzan et al. 2015). In this study, we apply one of the heat indices developed by Rothfusz (1990); it uses temperature and relative humidity and is applied primarily by the

National Weather Service in the United States. Rothfusz (1990) performed a multiple regression analysis on the original table of heat indices computed by Steadman (1979). However, Rothfusz's (1990) equation is not applicable for all ranges of relative humidity and temperatures. An adjustment of the heat index equation is needed (http://www.wpc.ncep.noaa.gov/html/heatindex_equation.shtml). Using Rothfusz's (1990) equation and the adjustment mentioned above, we can compare how well the models capture the right tail of the PDF for heat index. The heat index value for each location is calculated for both the simulations and the observations, so the difference can be plotted on a grid. The maximum temperatures and near-surface relative humidity values are used to calculate the heat index. Maximum temperature is used instead of mean temperature to determine whether the errors in heat index are result from the known extreme maximum temperature biases discussed for Figure 12, or if relative humidity biases could affect the results more significantly in some regions.

Figure 14 shows the difference in heat index for each simulation's 95% threshold compared to the observations. Generally, the WRF_NCEP shows the smallest bias over the entire CONUS, followed by WRF_CCSM_nBC. There is large positive bias for heat index in the Southern Great Plains and the western CONUS for WRF_GFDLNN, WRF_GFDLN, and WRF_CCSM_BC that is not evident in the maximum temperature, which indicates that the relative humidity is overestimated in those regions. In comparison with WRF_GFDL_NN, nudging reduces the bias for heat index in the Northwest. In the Southeast, where heat index values tend to be the highest during the summer, the models without bias correction underestimate the 95% threshold. In contrast, the models that use bias correction slightly overestimate heat index and perform better over the Southeast. Overall, there are more differences between the two GFDL runs when nudging is applied and the two CCSM4 runs when bias correction is used. When comparing the GFDL and CCSM4 runs that make use of both bias correction and spectral nudging, there are still a couple of important differences. For example, in parts of the Midwest and central Plains, the differences in heat index threshold are as high ~6°C in some locations, indicating the biases in the boundary conditions are still significant in those areas.

4.3.3 Extreme Precipitation

Figure 15 shows the difference between the model and the observed 95% threshold for daily precipitation. All of the precipitation data is filtered to only include precipitation days that record at least 0.01 inches (or 0.254 mm), to guard against minimum unrealistic values. In comparison with GCM- or reanalysis-driven WRF runs, the mean and median of the six WRF simulations show a significant dry bias in extreme precipitation. WRF_NCEP shows a wet bias over the Great Plains for not only the daily mean precipitation (as shown in the top left panel of Figure 10), but also the extreme precipitation. This may be due to the short spin-up time and/or the strong nudging strength applied in this simulation (WK14), which has been modified in the GCM-driven runs. The five GCM-driven simulations underestimate the extreme precipitation by 3.5–8.3 mm over the southeast. This is likely because these simulations lack the ability to capture regularly occurring small-scale convection in this region that cannot be fully resolved with 12-km horizontal resolution. In comparison to WRF_GFDLNN, WRF_GFDLN significantly

reduces the model bias in extreme precipitation over the Great Plains and the Midwest. In comparison to WRF_CCSM_nBC, WRF_CCSM_BC significantly reduces the bias in extreme precipitation over all regions except the Northwest and the Northeast. The WRF_HadGEM run, which does not use bias correction or spectral nudging, shows a much larger dry bias over most of the regions with the exception of the Northeast.

Extreme precipitation events occur frequently when daily precipitation values are to the right of the 95% threshold on the PDF curve for multiple consecutive days (Janssen et al. 2013). In many cases, the heaviest precipitation events occur because a storm system is stagnant over similar areas on consecutive days (e.g., Francis and Vavrus 2012). Although many other environmental factors determine the extent and magnitude of flash floods (Montz and Gruntfest 2002), the best these models can do is attempt to improve on the forecasting frequency of long-term extreme precipitation events. For this reason, in addition to daily precipitation extremes, this study also analyzes the model's ability to simulate major precipitation events for 2- and 3-day storm totals. Figure 16 shows the differences in frequency of the 99% threshold for 2-consecutive-day precipitation. By finding the 99% average regional threshold for 2-day precipitation events from the observations, the difference in the number of times the model predicts this occurrence shows how well the simulation handles storm system movement across the United States. This is calculated by ranking all of the total 2-day precipitation events at each location that experienced at least a trace of precipitation, and then calculating the number of occurrences that are greater than the regional observed threshold for the whole decade in each of the six simulations. Figure 16 shows the number of times the model output was greater than the regionally averaged 99% threshold in the reference data and is standardized by subtracting the number of events in the reference data at each grid point that were greater than the 99% threshold. The reason the difference is calculated is that, depending on the location or region, there may be a high frequency of precipitation days, which means that more 99% events would be expected for these locations over the course of a decade. The 2-day and 3-day results for this metric are similar enough that we only present the difference in 2-day precipitation extremes in this study.

The GCM-driven simulations tend to underestimate the frequency of 99% events along the Gulf of Mexico in the Southeast, and along the West Coast. Other regions, such as the Midwest, have differences in regional signs for each of the six simulations. When used together, bias correction and nudging tend to slow storm system movement across the Great Plains and the Midwest, as indicated by the positive 2-day precipitation anomalies. Without nudging or bias correcting the boundary conditions, the WRF simulations move storm systems across the central United States faster, which leads to fewer events that meet the observed 99% threshold criteria for that location. The addition of nudging in the GFDL runs enhances a strong positive bias in a large area of the Southwest—as well as through most of the Northern Plain states—that is not present in the no-nudging run. The WRF_GFDLN run reduces high negative bias in the WRF_GFDLNN in much of the Midwest. To a lesser extent, bias correction also reduces this same negative anomaly for the CCSM runs in most of the Midwest.

4.3.4 Modified Algorithm for Estimating Extremes in a Distribution, Generalized Extreme Value Theory

We applied the generalized extreme value (GEV) distribution, which unites the Gumbel, Fréchet, and Weibull distributions into a single family to allow for a continuous range of possible shapes (Coles 2001) and has been applied widely in studies of extreme climate. The GEV distribution adopts a "divide and conquer" strategy, breaking the variable into several aspects conceptually to reduce uncertainties in assessing RCM performance (Katz et al. 2013). The GEV distribution is a powerful approach for estimating the extremes (or return level) with given return periods and predicting extreme changes in the future. The three parameters for the GEV model are location (where the maxima are concentrated), scale (the spread of the distribution around the median), and shape (shown in Figure 17).

We developed a "borrowing strength" GEV model with two important assumptions. First, given the climate, our GEV model assumes that the shape parameters within an area 96 km × 96 km (eight grid points for 12-km WRF simulation; three grid points for NARR) are constant. Given this area, the scale parameters and location parameters are dependent on each other. The parameters for different areas, however, are independent. In addition, we assume a spatiotemporal model for the location parameter—with a spatially varying intercept term and a

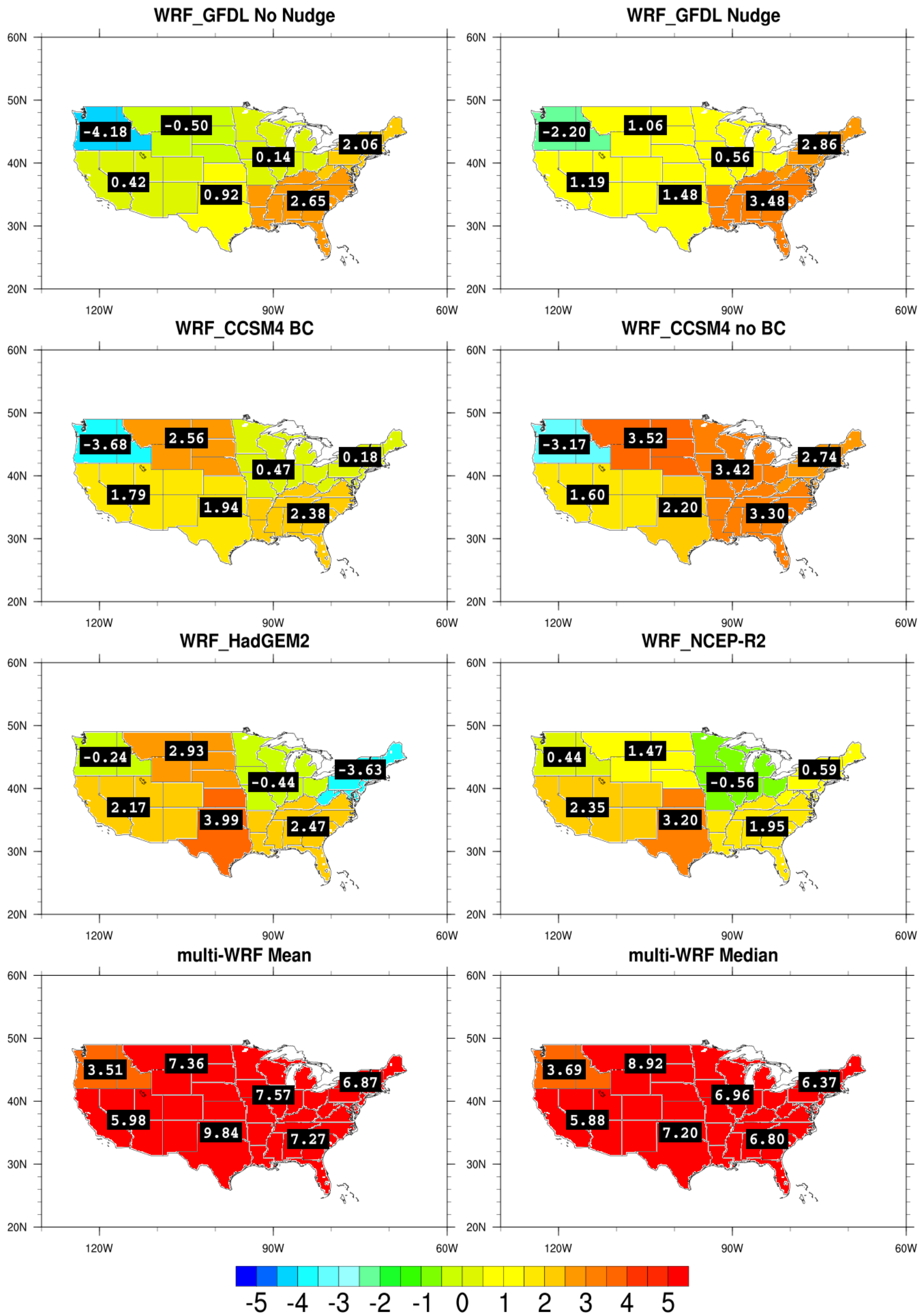

FIGURE 13 Average Regional Differences in 5% DJF Minimum Temperature Threshold Events between the Models and Observations (values for each region are in °C)

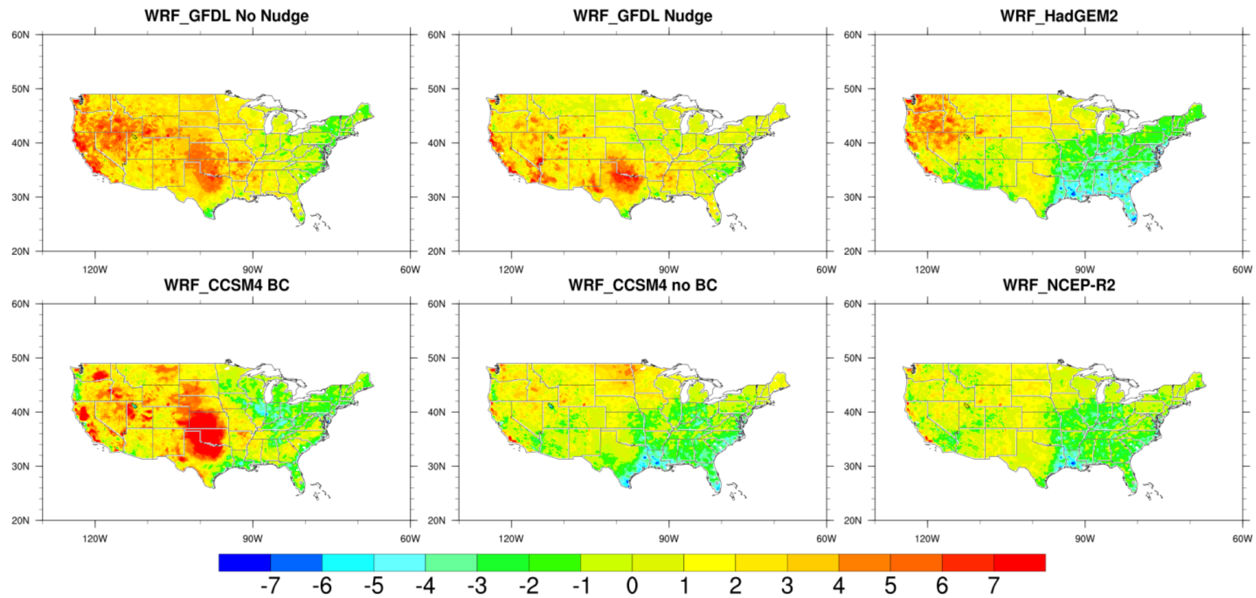

FIGURE 14 Difference in 95% Heat Index Threshold (in °C) between the Six Model Simulations and Observation

spatially varying long-term trend that was shown to be better at describing the data in a warming climate than the same model with a time-invariant location parameter (Brown et al. 2008; Craigmile and Guttorp 2013). This approach was applied to evaluate 31-year WRF-downscaled extreme maximum temperature through comparison with NARR data. Uncertainty in GEV parameter estimates and the statistical significance in the differences of estimates between WRF and NARR are accounted for by conducting a novel bootstrap procedure that makes no assumption of temporal or spatial independence within a year, which is especially important for climate data. Despite certain biases over parts of the United States, overall, WRF shows good agreement with NARR in the spatial pattern and magnitude of GEV parameter estimates. Both WRF and NARR show a significant increase in extreme maximum temperature over the southern Great Plains and southeastern United States in January and over the western United States in July. The GEV model shows clear benefits from the regionally constant shape parameter assumption, for example, leading to estimates of the location and scale parameters of the model that show coherent spatial patterns.

4.4 MODEL UNCERTAINTY ANALYSIS

As we stated in Section 2, the three primary causes of model uncertainty are usually model parametric uncertainty, model internal variability, and scenario uncertainty. Because we have only a few ensemble members, we could not analyze the scenario uncertainty fully; however, we will combine our analysis with the models used for simulating for the NARCCAP project to develop a weighted-mean uncertainty. We analyzed the model sensitivities to different physics schemes and model setup, as well as the model internal variabilities due to different initializations.

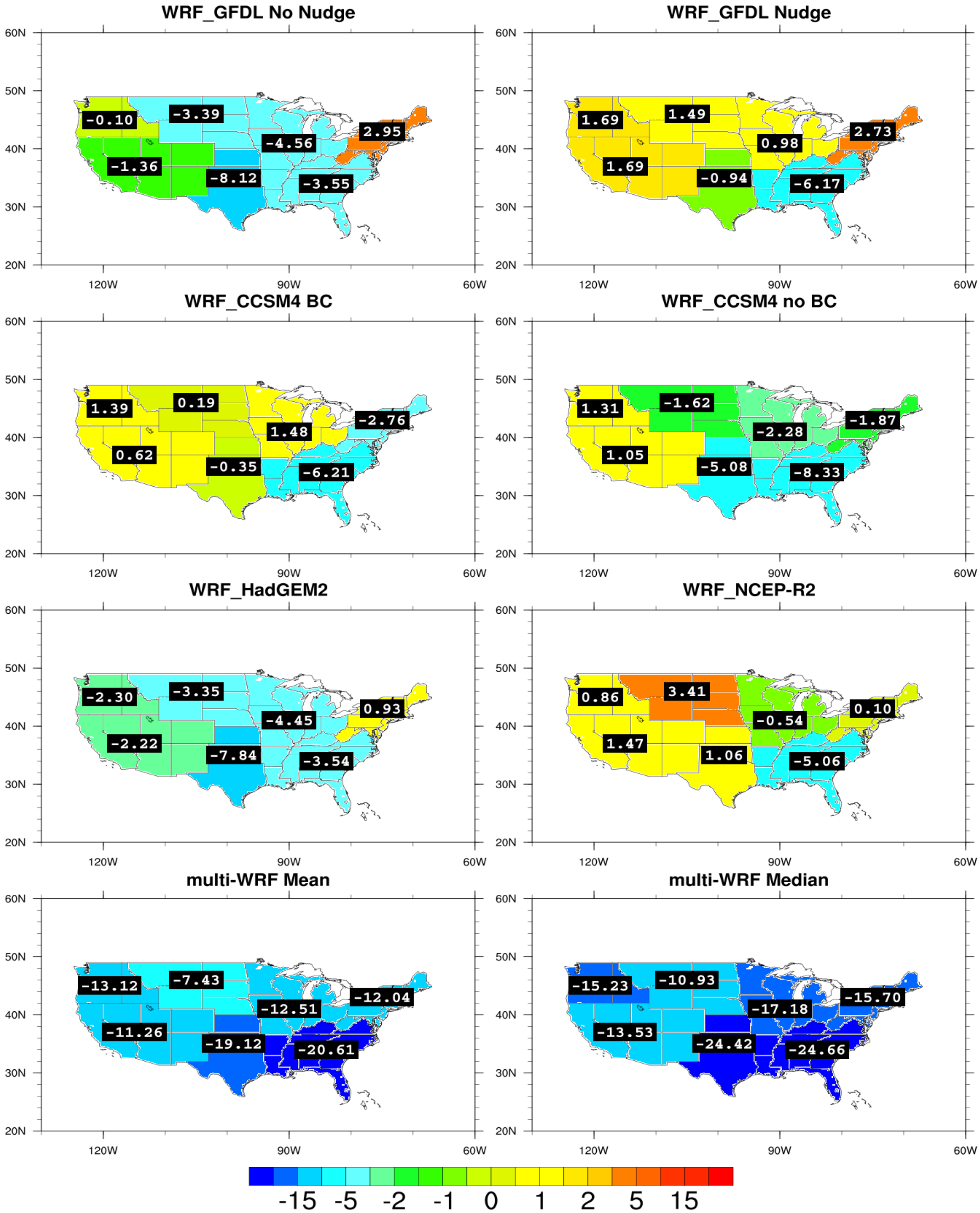

FIGURE 15 Average Regional Difference in 95% Threshold Extreme Precipitation Events between the Models and Observations

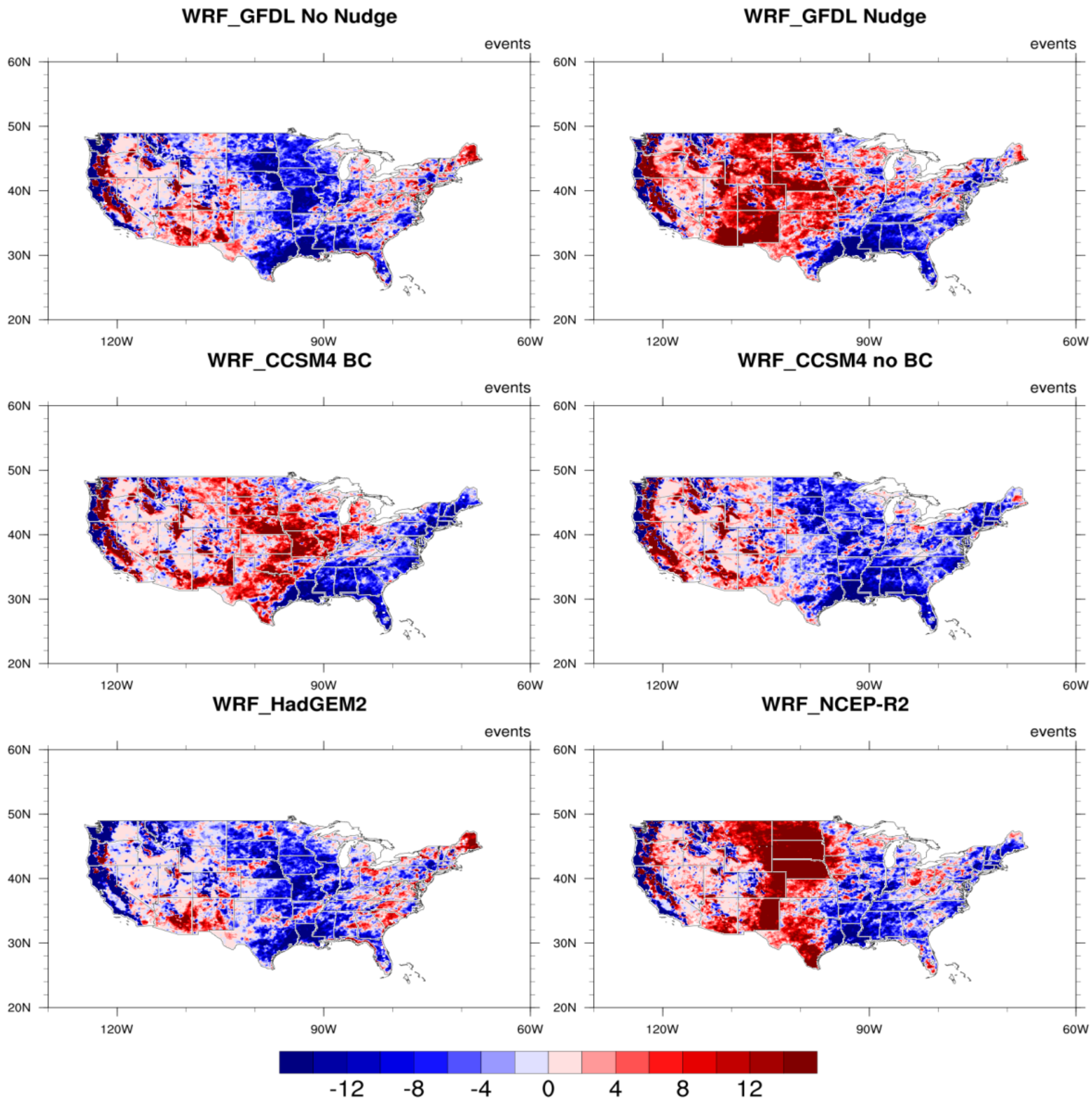

FIGURE 16 Differences in the Frequency of 99% Threshold Events between Models and Observations for 2-day Precipitation Events (In order to be categorized as an "event," the grid point must experience at least a trace of precipitation for 2 consecutive days. To standardize these values, the difference between the number of 99% events in the observations is subtracted from the model values.)

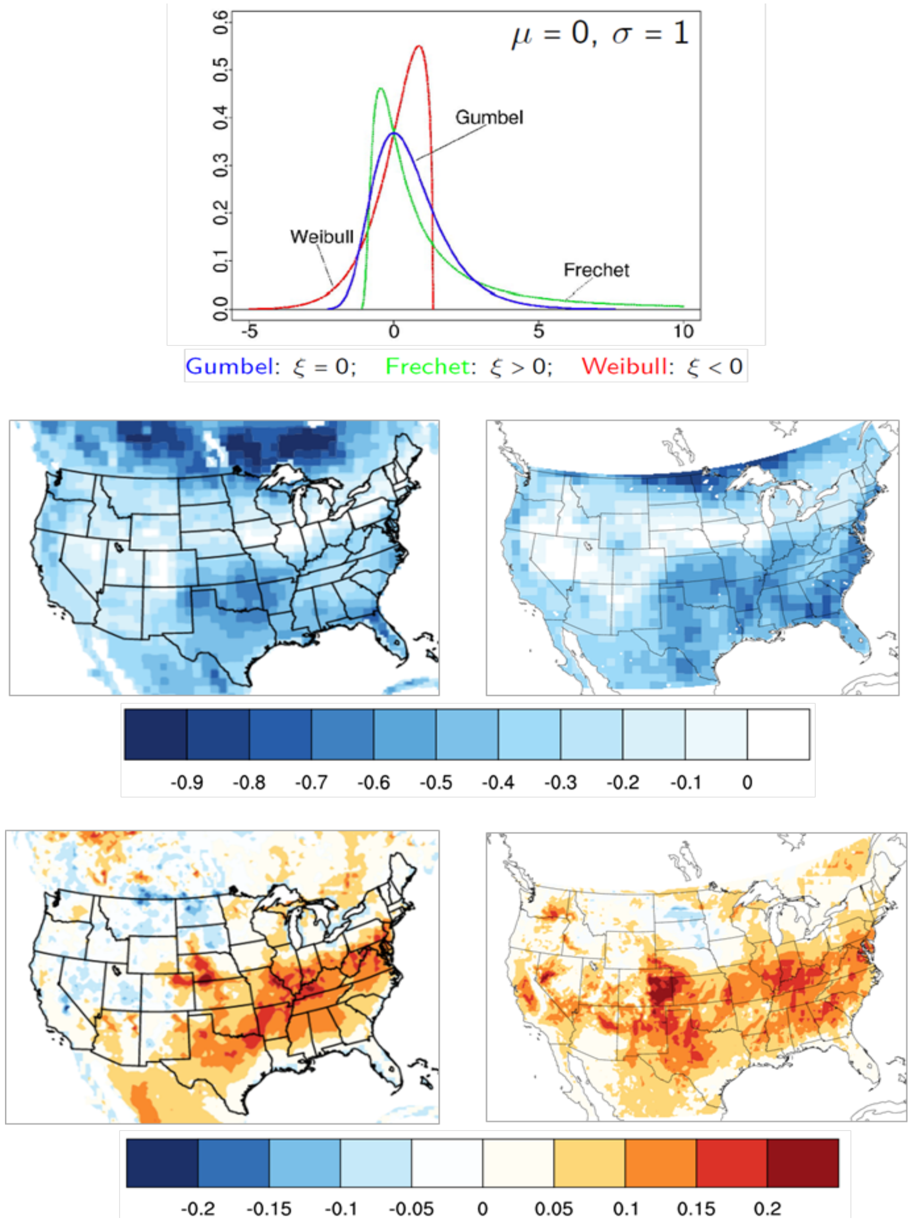

FIGURE 17 Top: Three Types of Distributions Used in the GEV Model Applied for Estimating Extremes and Repeat Periods in This Study (extreme temperature usually follows the Weibull distribution, while extreme precipitation usually follows the Fréchet distribution); Middle: WRF simulation Capturing the Shape of the Distribution Well (left: NARR; right: WRF); Bottom: Long-term Trend of January Extreme Maximum Temperature, with a Positive Trend (0.1–0.2 K/30 years) for January Extreme Maximum Temperature, and the WRF Model Reasonably Capturing the Magnitude of the Trend with Bias Over the Southwestern United States

4.4.1 Model Sensitivities to Physics Parameterizations

To explore the potential reasons for model biases, we first compared the diurnal variations in precipitation between the NARR values and the WRF results over the CONUS. The diurnal cycle of precipitation is dominant over the sub-synoptic and synoptic cycles for summer precipitation over most of the CONUS, as shown by Castro et al. (2007) on the basis of spectral analysis of integrated moisture flux convergence. Then we conducted several experiments to test the model sensitivities to convective parameterization, microphysics scheme, spectral nudging strength, and spin-up time. In our historical simulation, the most significant problem with the model calculations was the wet bias over the Great Plains in the warm season and the warm bias over the South in all four seasons. A wet bias over the mountain ranges in cold seasons in the WRF calculations showed pronounced improvement compared to the NARCCAP-WRFG results.

Our sensitivity experiments (Figure 18 and Tables 3 and 4) showed that spectral nudging strength, spin-up time, integration method (re-initialization versus continuous integration), and microphysics scheme had important effects on the calculated precipitation and temperature, but the cumulus parameterizations tested showed no effect. Reducing the spectral nudging strength and/or allowing longer spin-up time can partly address the wet bias over the Great Plains and the Desert and the warm bias over the South, but these changes generated larger bias for precipitation over the South and the Rockies, and for temperature over the Great Plains and the Desert. Replacing the WSM6 microphysics scheme with the Morrison scheme reduced the wet bias over the western mountain ranges and over the Great Plains in winter. Turning off the spectral nudging also reduced the excessive rainfall over the Great Plains in summer but generated larger bias for precipitation over the southeastern and eastern CONUS (Wang and Kotamarthi 2013). The biases in the WRF simulation could also result from uncertainties in the LBCs. Liang et al. (2004) found that using the European Center for Medium-Range Weather Forecast reanalysis can address the winter dry biases over the South, which likely result from LBC errors for NCEP-R2 data.

4.4.2 Internal Variability

We investigated the internal variability (IV) of a regional climate model, considering the impacts of horizontal resolution and spectral nudging. Ten-member ensembles of 1-year simulations using the WRF model are conducted for three sets of model configurations. This includes simulations at spatial resolutions of 50 km and 12 km without spectral nudging and simulations at a spatial resolution of 12 km with spectral nudging. All of those simulations cover an entire annual cycle and are generated over the same domain—much of North America. The degree of IV is measured as the spread between the individual members of the ensemble during the integration period. The IV is defined by the spread among the ensemble members during the

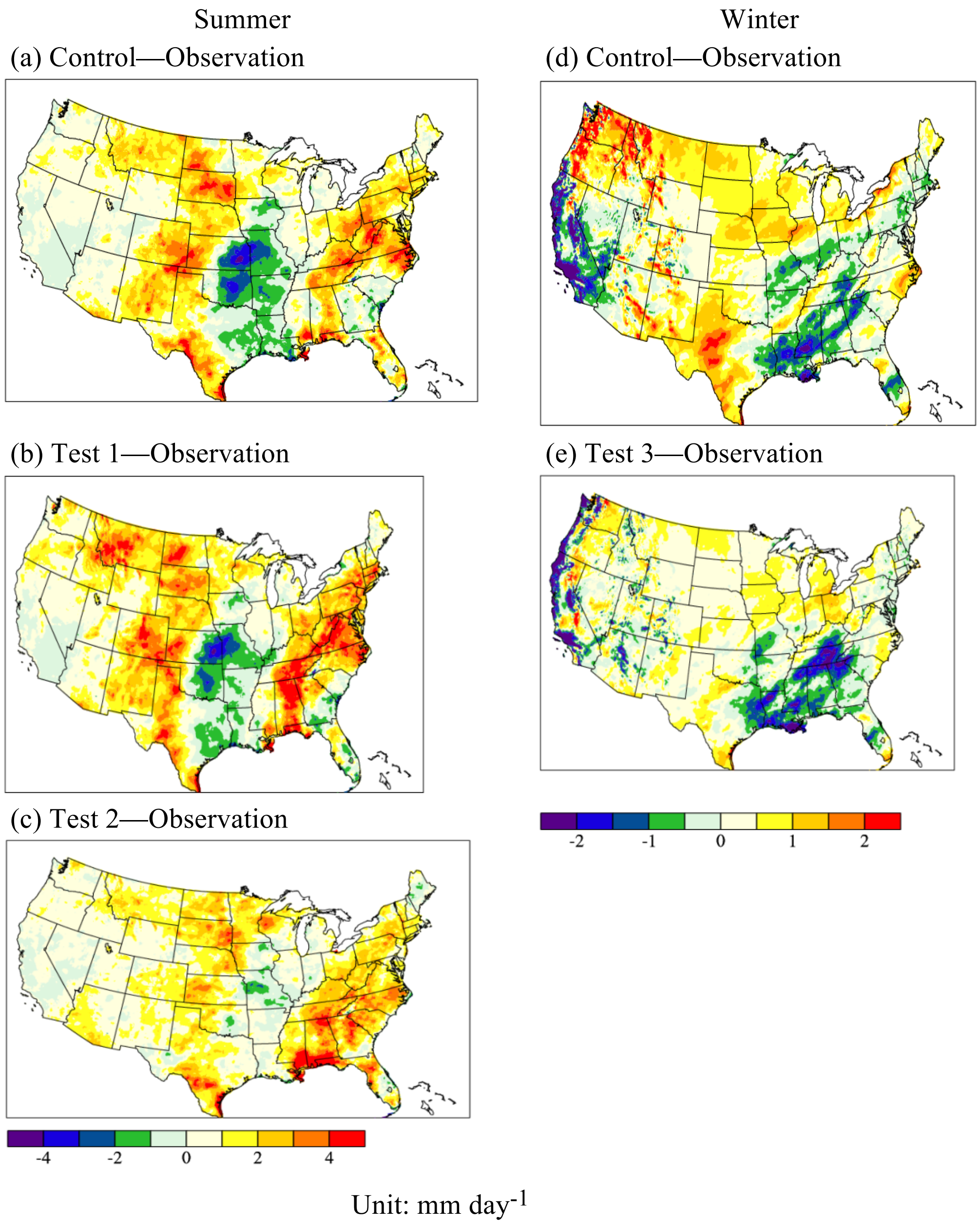

FIGURE 18 Precipitation in Summer (left) and Winter (right) 2005 for the Difference between the (a) Control Simulation and (d) PRISM Observation, (b) Difference between Test 1 and PRISM Observation, (c) Difference between Test 2 and PRISM Observation, and (e) Difference between Test 3 and PRISM Observation (Test 1 replaces the Grell-Devenyi cumulus parameterization with Kain-Fritsch; Test 2 reduces the spectral nudging from 3×10^{-4} s^{-1} to 3×10^{-5} s^{-1}. Test 3 replaces the WSM6 microphysics with Morrison.)

TABLE 3 Control- and Test 4–simulated Regional Average Biases for Temperature in 2005[a,b]

	Temperature Bias (°C)							
	GP		Southwest		SC		NR	
Season	Control	Test 4	Control	Test 4	Control	Test 4	Control	Test 4
Spring	1.98	0.45	0.05	**0.08**	2.16	1.77	0.34	0.31
Summer	1.38	**1.61**	0.71	**1.96**	0.61	0.37	-0.83	0.10
Fall	2.04	**2.94**	0.37	**1.77**	2.35	**2.47**	0.93	**1.66**
Winter	1.54	**2.78**	-0.59	**1.11**	3.92	3.82	2.16	**4.23**

[a] Test 4 runs a 2-year continuous simulation (2004–2005) with 1-year (2004) spin up.
[b] Bold font indicates a larger bias for Test 4 than for the control simulations.
Note: GP is Great Plains; SC – South Central USA; NR is Nortern Rockies

TABLE 4 Control- and Test 4–simulated Regional Average Biases for Precipitation in 2005[a,b]

	Precipitation Bias (mm day^{-1})							
	GP		Southwest		SC		NR	
Season	Control	Test 4	Control	Test 4	Control	Test 4	Control	Test 4
Spring	1.44	1.02	3.65	0.70	1.61	1.08	1.55	**1.70**
Summer	1.14	0.07	4.20	1.58	1.70	**3.13**	0.51	0.35
Fall	0.57	0.18	1.23	0.27	-0.12	**-0.58**	0.89	**1.28**
Winter	0.62	**0.77**	1.50	0.28	-0.11	**-0.43**	1.45	**2.61**

[a] Test 4 runs a 2-year continuous simulation (2004–2005) with 1-year (2004) spin up.
[b] Bold font indicates a larger bias for Test 4 than for the control simulations.

integration period. The spread is measured by the standard deviation between the 10 members in the ensemble. First, we calculate the variance of the 10 members:

$$\sigma_{en}^2(i,j,t) = \frac{1}{N}\sum_{n=1}^{N}\left[Y_n(i,j,t) - \overline{Y}(i,j,t)\right]^2 \tag{1}$$

Where $Y_n(i, j, t)$ refers to a variable Y on grid point (i, j) at time t for member n in the ensemble and N is the total number of ensemble members, here $N = 10$. $\overline{Y}(i,j,t)$ is the ensemble mean defined by eq. (2):

$$\overline{Y}(i,j,t) = \frac{1}{N}\sum_{n=1}^{N} Y_n(i,j,t) \tag{2}$$

A measure of the seasonal average of the IV and its geographical distribution over the model domain is calculated by the square root of the seasonal average of $\sigma_{en}^2(i,j,t)$ in eq. (1). Domain-averaged IV during the course of the model integration is calculated by the square root of the spatial average of $\sigma_{en}^2(i,j,t)$ in eq. (1). The details of these calculations can be found in Alexandru et al. (2007) and Lucas-Picher et al. (2008). The IV at a spatial resolution of 12 km with spectral nudging is also compared with the climate change signal projected by the same model configuration. The variables investigated are precipitation, near-surface air temperature, relative humidity, wind, and sea level pressure, as well as geopotential height at 500 hPa. We focus here on precipitation. Figure 19 shows the IV of precipitation in four seasons produced by 50km_no_nudg, 12km_no_nudg, and 12km_nudg. There is a clear seasonal cycle for the IV of each set of simulations, with the largest IV in summer and the smallest IV in winter. It is interesting that the geographic distributions of the precipitation IV generated by nudged runs (12km_nudg) and non-nudged runs (50km_no_nudg and 12km_no_nudg) are different. Non-nudged runs show relatively large IVs over eastern North America for all four seasons, especially over the southeastern CONUS in summer, where large convective precipitation occurs. In contrast, they show relatively small IVs over western North America, especially in the northwestern part of the domain. Alexandru et al. (2007) also found large precipitation IVs in the southeastern CONUS in June through August by running a regional climate model without using spectral nudging. They suggested that there is a link between 850-hPa geopotential height in the Northeast and precipitation in the Southeast. The precipitation is a triggering mechanism for the geopotential height, which continues to develop along the storm track and reaches its maximum toward the Northeast. The 500-hPa geopotential height in this study shows the same geographic distribution, with the largest IVs over the Northeast. Comparing the three sets of simulations, 12km_no_nudg shows the largest IV, while 12km_nudg shows the smallest IV. These differences can be seen in all four seasons.

Overall, the results show that, although the IVs exhibit a clear annual cycle with larger values in summer and smaller values in winter, they are smaller at the spatial resolution of 50 km than at 12 km when nudging is not applied. Applying nudging to simulations at 12 km reduces the IV by a factor of 2, and produces smaller IVs than the simulations at 50 km without nudging. Applying nudging also changes the geographic distributions of IV in all the examined variables. The IV is much smaller than the inter-annual variability at the seasonal scale for regional averages of temperature and precipitation. The IV is also smaller than the climate change signal of temperature in mid- and late 21st century. However, the uncertainty due to IV plays an important role in the climate change signal of precipitation, especially in summer and fall and in the mid-21st century.

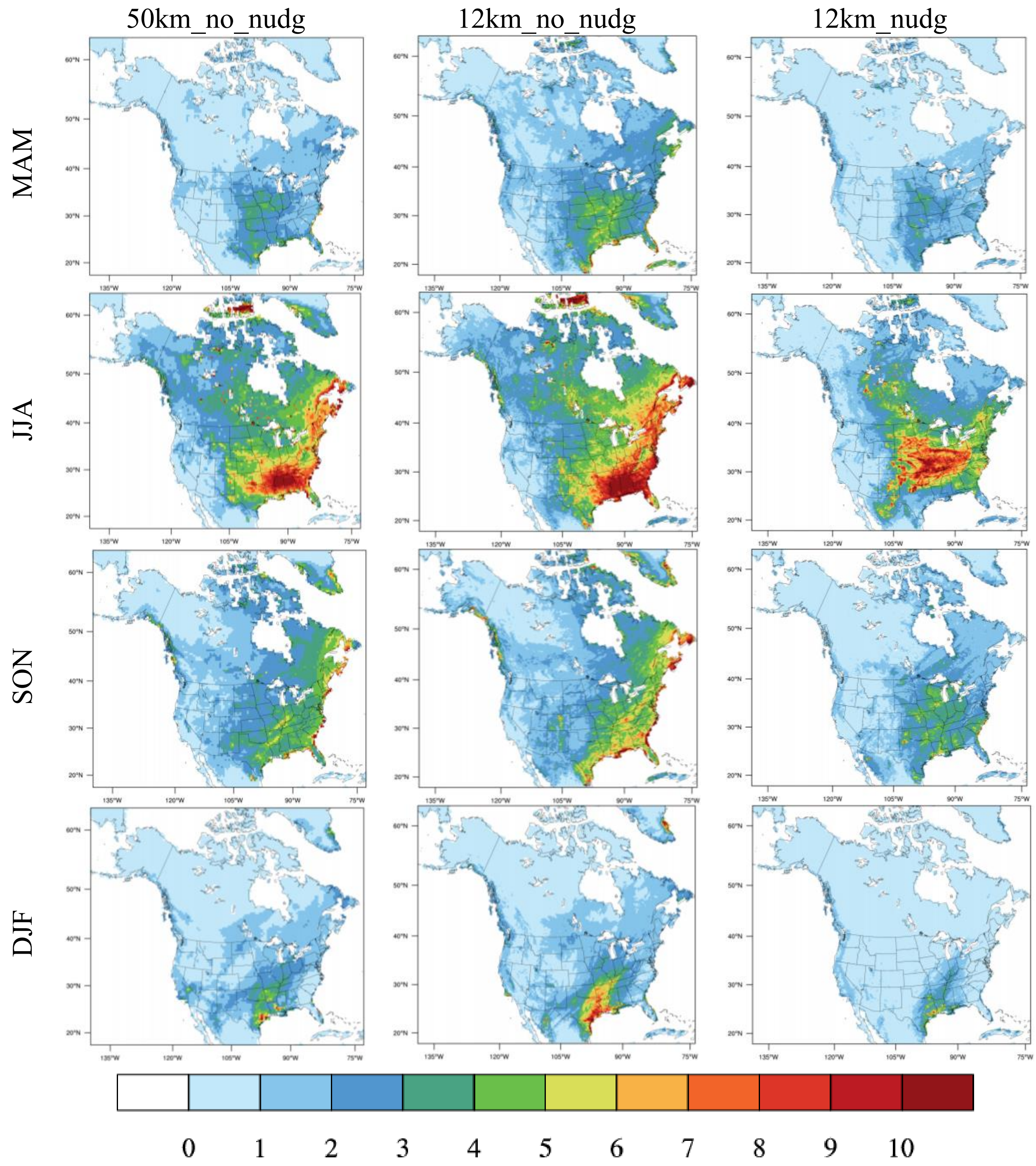

FIGURE 19 Geographic Distribution of Internal variability of Precipitation Amount (mm/day) in Four Seasons for the Year 1995; from Left to Right, the Panels Show Simulations on a 50-km Resolution Grid with No Spectral Nudging, 12-km Spatial Resolution Model with No Nudging, and 12-km Spatial Resolution Model with Spectral Nudging

5 FUTURE PROJECTIONS

In this section we present results from the projections made using the model for the period 2045–2054 and 2085–2094 under two greenhouse gas (GHG) forcing scenarios, RCP 4.5 and RCP 8.5, respectively.

5.1 BIAS CORRECTION FOR RCMs AND GCMs

This study applies the approach tested by Bruyère et al. (2013), which corrects the mean errors in the GCM but retains the GCM's 6-hour weather, longer-period climate variability, and climate change (Figure 1 in Xu and Yang 2012; Figure 4 in Bruyère et al. 2013). Thus, it allows the variance, diurnal cycle, seasonal cycle, and phase of interannual variations to change freely from the past to future periods. In this study, we correct the atmospheric components of CCSM4 (used as boundary conditions for the downscaling) by using the National Center for Atmospheric Research Reanalysis Project (NNRP) data (Kalnay et al. 1996) over the period 1950–1979, which we define as the base period. This period was chosen because it has no significant climate trend, and thus no trend needs to be removed before bias correction. The process used is described below.

First, the 6-hour NNRP data and CCSM4 outputs are broken down into a climatological mean plus a perturbation term:

$$\begin{aligned} CCSM &= \overline{CCSM} + CCSM' \\ NNRP &= \overline{NNRP} + NNRP' \end{aligned}$$

Accordingly, the CCSM4 model output for the three time periods we model (1994–2004, 2044–2054, and 2084–2094) can be written as follows:

$$\begin{aligned} CCSM_c &= \overline{CCSM_c} + CCSM_c' \\ &= \left(\overline{CCSM_b} - \overline{NNRP_b}\right) + \overline{NNRP_b} + \left(\overline{CCSM_c} - \overline{CCSM_b}\right) + CCSM_c' \end{aligned}$$

The subscripts *b* and *c* represent the base period (1950–1979) and the considered period (1994–2004, 2044–2054, or 2084–2094), respectively. Thus, the bias-corrected CCSM4 data $CCSM_c^*$ in the three considered periods are constructed by removing the CCSM4's climatological bias $\overline{CCSM_b} - \overline{NNRP_b}$:

$$\begin{aligned} CCSM_c^* &= \overline{NNRP_b} + \left(\overline{CCSM_c} - \overline{CCSM_b}\right) + CCSM_c' \\ &= \left(\overline{NNRP_b} - \overline{CCSM_b}\right) + \overline{CCSM_c} + CCSM_c' \\ &= \left(\overline{NNRP_b} - \overline{CCSM_b}\right) + CCSM_c \end{aligned}$$

The corrected atmospheric variables include zonal and meridional wind, geopotential height, temperature, and relative humidity every 6 hours, for three dimensions. The bias in sea surface temperature (SST) is corrected by using Analysis SST data, which merges Hadley Centre and National Oceanic and Atmospheric Administration optimum interpolation SST datasets (Hurrell et al. 2008) over the period 1950–1979. These variables were tested by Bruyère et al. (2013) and found to be the most important for GCM bias corrections. In addition, we correct the land/sea mask in the land surface model by replacing “land” with “sea” over the Great Lakes region, as first recommended by Gao et al. (2012) in a sensitivity study showing that modification of the land/sea mask could significantly reduce the bias of the 2-m air temperature near the Great Lakes simulated by the WRF model.

5.2 CLIMATOLOGICAL MEANS—PRECIPITATION

Here we present result from the WRF model version 3.3.1, used to dynamically downscale CCSM4, in one historical period (1995–2004) and two future periods (2045–2054 and 2085–2094) under RCP 4.5 and RCP 8.5. The WRF model is applied at a horizontal resolution of 12 km, with 600 west–east × 515 south–north grid points and 28 vertical levels over most of North America (Figure 2a). As discussed earlier, the bias-correction approach applied in this study only corrects the climatological mean of CCSM4 and allows the data ($CCSM_c^*$) to change freely at subdaily, daily, seasonal, and yearly scales. Therefore, the original CCSM4 and corrected CCSM4 have the same variabilities for those corrected atmospheric variables. The result is mostly similar changes in precipitation from the historical period to the future, as projected by BC_WRF (bias-corrected WRF) and No_BC_WRF (WRF simulation with no bias correction), although the absolute values for future precipitation projected by BC_WRF and No_BC_WRF are different. We present the future changes projected by BC_WRF in this section. The changes projected by No_BC_WRF are similar in geographic patterns and magnitudes, with differences in magnitude less than 10%.

Figures 20 and 21 show the annual and seasonal mean precipitation changes in the mid- and late 21st century (2045–2054 minus 1995–2004 and 2085–2094 minus 1995–2004) under RCP 4.5 and 8.5, as projected by BC_WRF and CCSM4. The precipitation change signals in the late 21st century, especially the wet tendency over Canada and Alaska, are generally stronger for both WRF and CCSM4 than those in the mid-21st century under RCP 8.5. The changes are smaller under RCP 4.5 than under RCP 8.5. The annual means of changes in precipitation show decreases over the Desert and the south part of the Central region under both RCPs, as projected by WRF and CCSM4. WRF also shows increased precipitation over the eastern CONUS under RCP 8.5, while the WRF changes under RCP 4.5 and the CCSM4 changes under both RCPs are smaller.

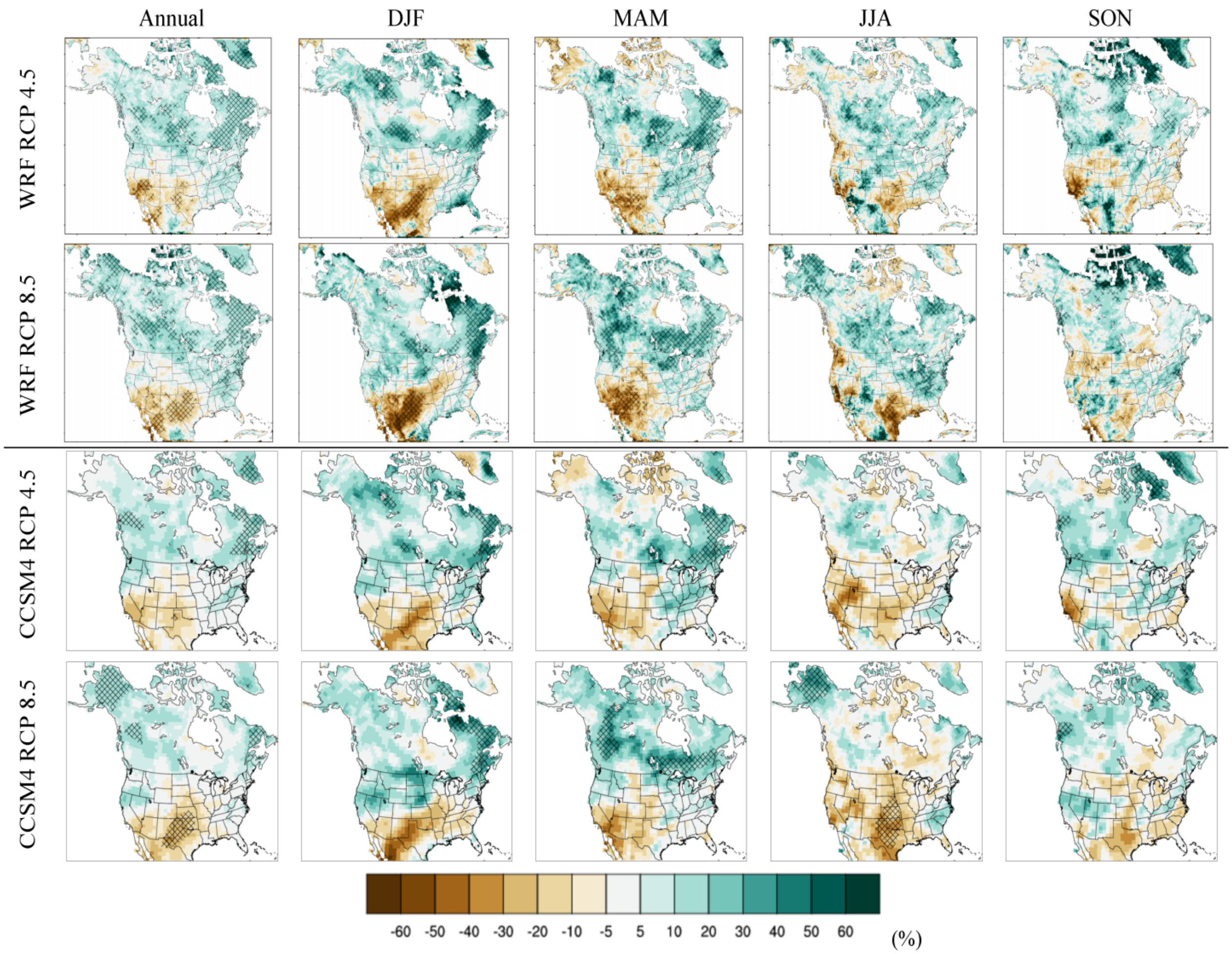

FIGURE 20 BC_WRF- and CCSM4-projected Change (%) in Seasonal and Annual Mean Precipitation for 2045–2054 versus 1995–2004 and for RCP 4.5 and RCP 8.5 (Cross-hatching indicates statistically significant changes. DJF = December, January, February. MAM = March, April, May. JJA = June, July, August. SON = September, October, November.)

The changes in precipitation show strong dependence on season. A strong decrease in precipitation is projected by both WRF and CCSM4 over the Desert and the southern part of the Central and the Mountain West in winter and spring. Similar results were found by previous studies using CMIP5 ensemble runs (Cook and Seager 2013) and downscaling simulations (Gao et al. 2012). In contrast, WRF projects increased precipitation in summer and fall over the southwestern CONUS under RCP 8.5 in 2045–2054, and the increase gains strength in 2085–2094 (Figure 9). On average, the precipitation intensity in summer over this region (mostly North American monsoon region) is projected to increase from 1.8 mm/day to 2.3 mm/day in 2045–2054 and to 2.7 mm/day in 2085–2094 under RCP 8.5. Cook and Seager (2013) also found increased precipitation over the Desert and the Southwest in September and October by using 41 CMIP5 ensemble members. Torres-Alavez et al. (2014) found increases in precipitation and related moisture flux in both summer and fall over the Desert and the Southwest in six CMIP5 model runs. On the other hand, slight increases in wet conditions over the eastern United States

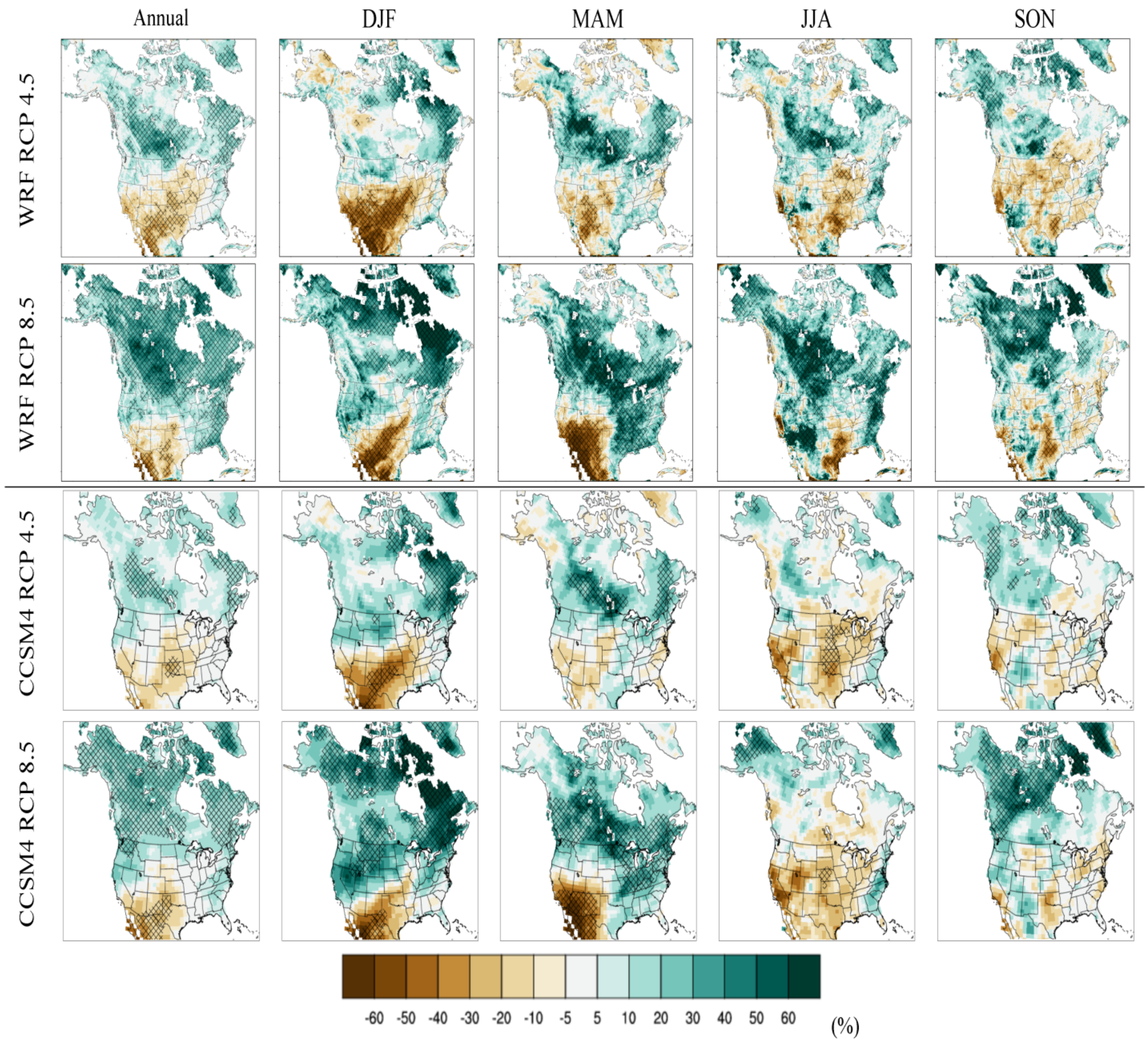

FIGURE 21 BC_WRF- and CCSM4-projected Change (%) in Seasonal and Annual Mean Precipitation for 2085–2094 versus 1995–2004 and for RCP 4.5 and RCP 8.5 (Cross-hatching indicates statistically significant changes. DJF = December, January, February. MAM = March, April, May. JJA = June, July, August. SON = September, October, November.)

are projected in spring by both WRF and CCSM4 and in summer by WRF. CCSM4 and WRF show similar patterns and signs (increase or decrease) for precipitation changes in winter, spring, and fall, but the magnitude is somewhat weaker in CCSM4 projections than in WRF projections. However, the changes in precipitation WRF projects for summer differ from those projected by CCSM4, and the differences under RCP 8.5 are significant.

Because physics representations in RCMs are different from those in GCMs (Han and Roads 2004; Liang et al. 2006), transformation from GCM to RCM results in redistribution of the influxes of mass, energy, and momentum in the RCM domain. This is a direct consequence

of different representations of key physical processes, especially land-atmosphere-ocean and convection-cloud-radiation interactions (e.g., Liang et al. 2004a,b). Han and Roads (2004) compared the performance of an RCM and a GCM and found that the large differences between the RCM and the GCM are mainly due to differences in model physics (such as the cumulus parameterization) rather than differences in grid resolution, especially when the sub-grid processes are important, as during summer. Liang et al. (2006), Pan et al. (2001, 2004), and Han and Roads (2004) compared the changes projected by an RCM and a GCM and found that patterns of precipitation change are significantly different in summer.

5.3 PRECIPITATION PERCENTILES AND HEAVY PRECIPITATION

To analyze the heavy precipitation statistics, we employ the geoclimatic subregions developed by Bukovsky (2011) for subregional evaluations (Figure 2b). We select the same 10 compound subregions over land (Figure 3) as did Martynov et al. (2013): Arcticland, Boreal, Central, Desert, East, Great Lakes (GLakes), Mountain West (MtWest), Northwestern Pacific (PacificNW), Southwestern Pacific (PacificSW), and South.

Figure 22 shows the changes in days per year (considering interannual variability) with specified precipitation amounts (1–10 mm, 10–20 mm, and 20–40 mm) for the 10-year historical period (1995–2004) compared to two 10-year periods in the future (2045–2054 and 2085–2094). As the precipitation threshold increases, fewer days in the historical and future periods experience these conditions, as expected. Although relatively light precipitation (1–10 mm) shows larger changes in days per year than does heavier precipitation (10–20 and 20–40 mm), the changes in days with light precipitation show larger variabilities due to interannual variations. For the 10-year-average changes, all subregions (except for Desert) show increases in days with all types of precipitation. Among the 10 subregions, the Pacific Northwest shows the largest change in days with 10–20 mm and 20–40 mm in the mid- and late 21st century under both RCPs; these occur mostly in fall and winter. East shows the second largest change in days of all types of precipitation, mostly in spring and summer. Considering the interannual variability of both the historical and future periods, we find a wide spread for changes in days with specified precipitation, especially in 2045–2054 under both RCPs. In 2085–2094, the spread of interannual variability is narrower, and days with 10–20 mm of precipitation show clear increases over much of North America.

Figure 23 shows the changes in frequency (occurrences per year) of 2-day duration 5-year return events and 2-day duration 10-year return events for the historical period (1995–2004) and compares them to future periods (2045–2054 and 2085–2094) obtained using Janssen et al.'s (2013) method. Duration refers to the number of days over which precipitation is accumulated, and return is an average of the number of years between events. Thus, the frequency of an event in a given time series depends only on the return time and the length of the time series. In the historical period, we have 365 events of 2-day duration and 5-year return and ~183 events of 2-day duration and 10-year return. To count the number of occurrences of these events in the future period, we first determine thresholds for a given duration and return from the historical simulations. We then apply those thresholds to the projections to identify the events that exceed the thresholds, and we compare the number of these events in future periods with the number in the historical period.

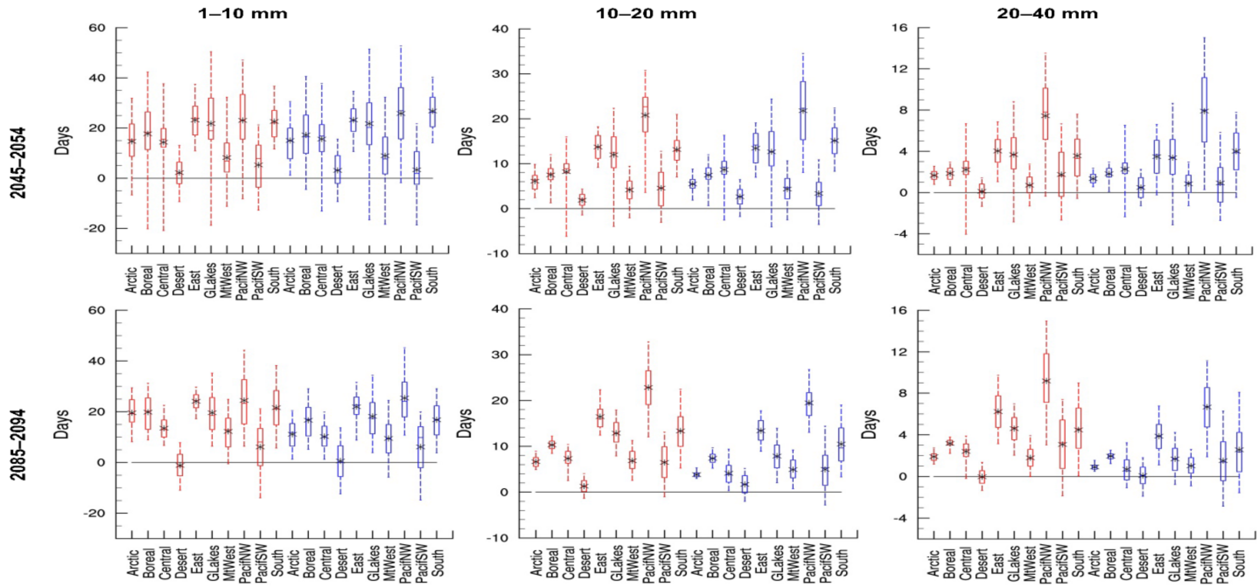

FIGURE 22 WRF-projected Changes in Days of Different Types of Precipitation for 1995–2004 versus 2045–2054 (top row) and for 1995–2004 versus 2085–2094 (bottom row) under RCP 8.5 (red boxes) and RCP 4.5 (blue boxes), Considering Interannual Variabilities (Boxes indicate the 25th and 75th quantiles, with the horizontal line indicating the median and the whiskers showing the extreme range of interannual variability. The stars indicate the 10-year average change between the historical and future periods.)

In 2085–2094, we find increases of 2-day events with 5-year and 10-year returns over much of North America except the Desert and Pacific Southwest, which show frequency decreases under both RCPs. The largest frequency increases are over the Arcticland and Boreal regions. In contrast, the Pacific Northwest and East (Figure 13) show the largest frequency increases of heavy precipitation. This indicates that long-duration events tend to increase more than short-duration and intense events over the Arcticland and Boreal regions. The 2-day duration events over the South show a slight decrease under RCP 4.5, but Figure 13 shows a significant increase for precipitation. This indicates that long-duration events tend to decrease, while the short-duration and intense events tend to increase in the future over the South. Over the Desert and Pacific Southwest, both types of extreme events are projected to be less frequent. We find similar changes in frequency of 2-day events with 5-year and 10-year returns over most of the ten subregions in the mid-21st century.

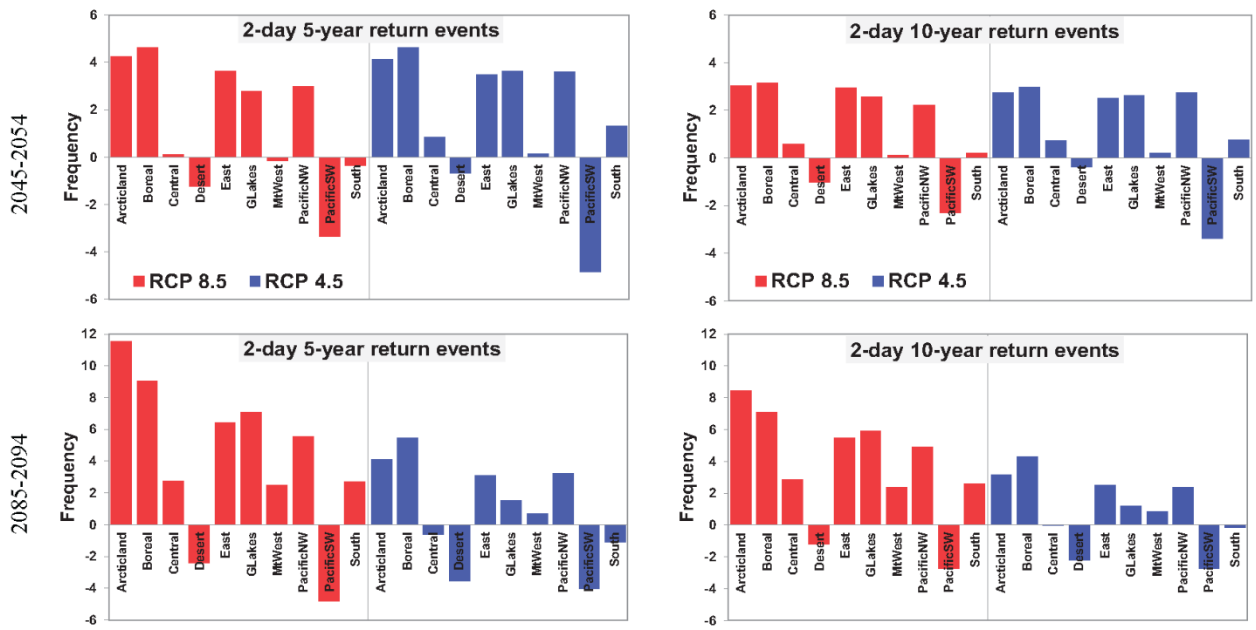

FIGURE 23 WRF-projected Changes in Frequency (number of occurrences per year) of 2-day Duration 5-year Return and 2-day Duration 10-year Return Events for 1995–2004 versus 2045–2054 (top row) and for 1995–2004 versus 2085–2094 (bottom row) under RCP 8.5 (red bars) and RCP 4.5 (blue bars)

5.4 CLIMATOLOGICAL MEANS—TEMPERATURE

The WRF model version 3.3.1, was used to dynamically downscale CCSM4, in one historical period (1995–2004) and two future periods (2045–2054 and 2085–2094) under RCP 4.5 and RCP 8.5 as discussed earlier. The WRF model is applied at a horizontal resolution of 12 km, with 600 west–east × 515 south–north grid points and 28 vertical levels over most of North America (Figure 3). As shown in Section 4.1, the bias-correction approach applied in this study only corrects the climatological mean of CCSM4 and allows the data ($CCSM_c^*$) to change freely at subdaily, daily, seasonal, and yearly scales. Therefore, the original CCSM4 and corrected CCSM4 have the same variabilities for those corrected atmospheric variables. We present the future changes projected by BC_WRF in this section for temperature.

The results presented below are for the 2085–2094 time period and for the two GHG forcing scenarios, RCP 4.5 and RCP 8.5. We first discuss the climatology of the projections of temperatures. Figure 24 shows the temperature change obtained for this time period compared to a decadal average for 1995–2004. The temperature during the winter months (left panel) is between 2 and 4°C over most of the continent. During the summer months we see changes on the order of 2 to 3°C. Winter changes are the largest over the northern part of the content in Alaska and Canada. The Southeast part of the CONUS sees the smallest changes.

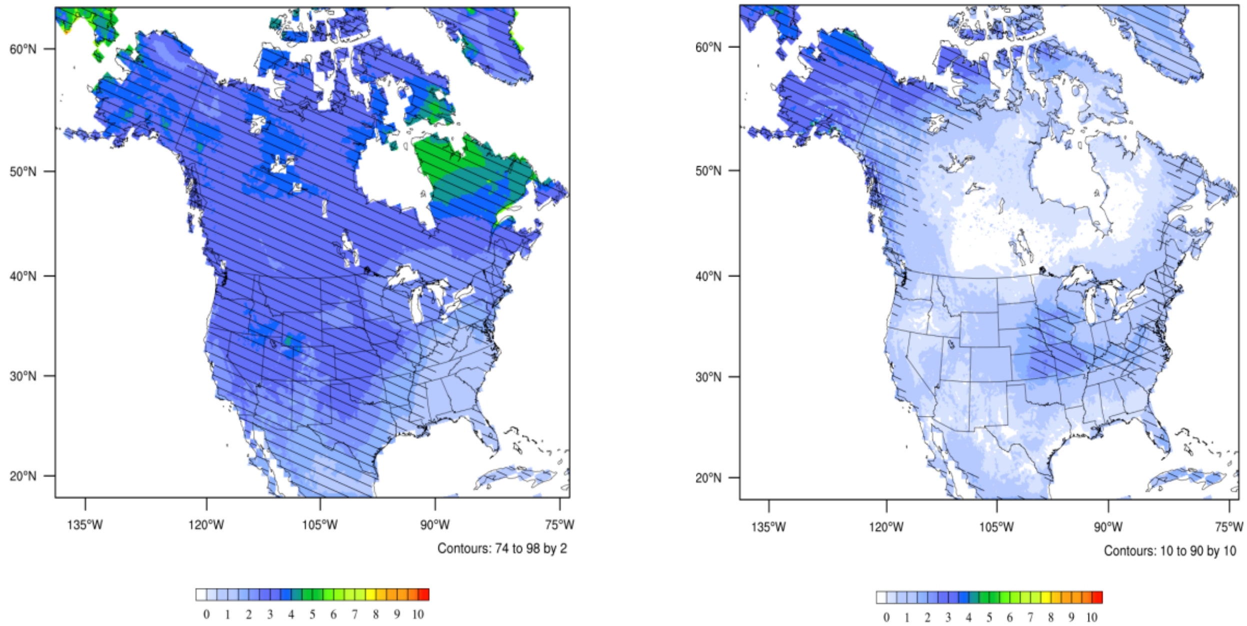

FIGURE 24 Projected Temperature Changes (in °C) for the Decade 2085–2094 Compared to 1995–2004 for the RCP 4.5 Scenario (The left panel is for the winter months and the right panel is for the summer months. The hatched lines indicate statistically significant changes based on t-test.)

In the summer months, the Midwest and the Pacific Northwest see the highest changes. Smaller changes are projected for central Canada. The hatching in Figure 25 indicates that the projected changes in temperature and statistically significant. The absence of hatching represents low confidence. For example, during the summer months the projected warming of 2°C in the Midwest has a high confidence, while that over the central parts of Canada has very low confidence.

A similar analysis was performed with RCP 8.5 scenario for the same time period (2085–2094), as shown in Figure 25. The changes again are presented compared to 1995–2004. The temperature changes at the end of the century compared to the end of the last century for winter months is projected to be 4 to 8°C higher over much of North America. The change projected for this higher GHG emission scenario is a nearly twice as large as the one shown for a mid-range GHG emission scenario (Figure 24). The highest winter temperature increase is again over the northern high latitudes with changes of nearly 7°C over the winter months compared to 1995–2004. Parts of the upper Midwest are projected to see temperature changes of nearly 5–6°C. The lowest temperature changes are along the Southeast's Atlantic coastal regions. The projections for the summer months show increases of over 4 to 6°C over most of the central portions of the CONUS. The lowest temperature increases are projected for the higher Northeast of the continent and along the coastal regions of the Southeast.

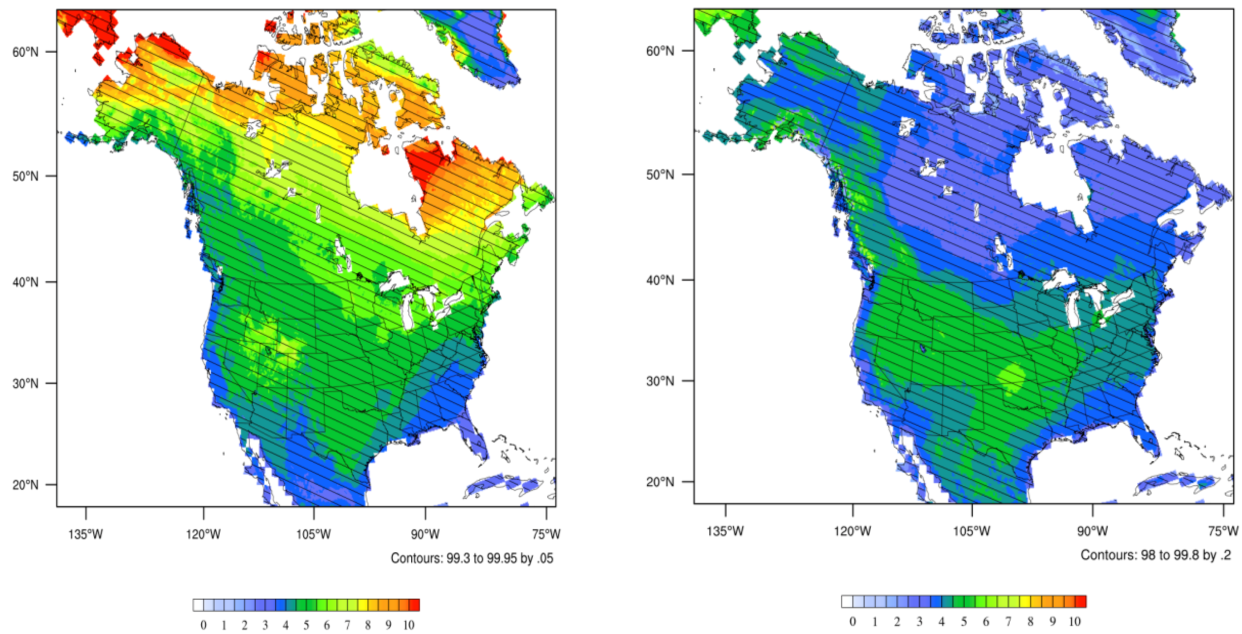

FIGURE 25 Projected Temperature Changes for the Decade 2085–2094 Compared to 1995–2004 for the RCP 8.5 Scenario (The left panel shows winter months and right panel shows summer months. The hatched lines indicate statistically significant changes based on t-tests.)

5.5 EXTREMES OF THE TEMPERATURE PROJECTIONS

The model results were analyzed to produce estimates of the extremes in temperature distribution for 2085–2094 as compared to 1995–2004 for the RCP 8.5 scenario. Figure 26 shows the increase in annual maximum (left panel) and annual minimum temperatures (right panel) recorded as a difference between these two decades. The annual maximum temperature increases by approximately 5°C over much of the CONUS and local regions in the Midwest experience a nearly 7°C increase in annual maximum temperatures. There is a widespread increase in the annual minimum temperature over the entire domain. Large portions of the continent's northern sections experience warming that falls between and 7 and 10°C in annual minimum temperature.

Figure 27 shows the increases in the number of days per year that are considered tropical nights (with daily minimum temperature greater than 20°C). A large portion of the CONUS is projected to experience an increase in number of days with high nighttime temperatures; this increase ranges between 40 and 80 days. The Southeast and the Eastern Seaboard experience the biggest changes, with increases on the order of 70 days or more. The right panel of Figure 27 shows the increase in the number of days per year that are considered summer days (with daily maximum temperature greater than 25°C). The increase is greater than 70 days in the central United States and between 20 and 40 days for the rest of the CONUS.

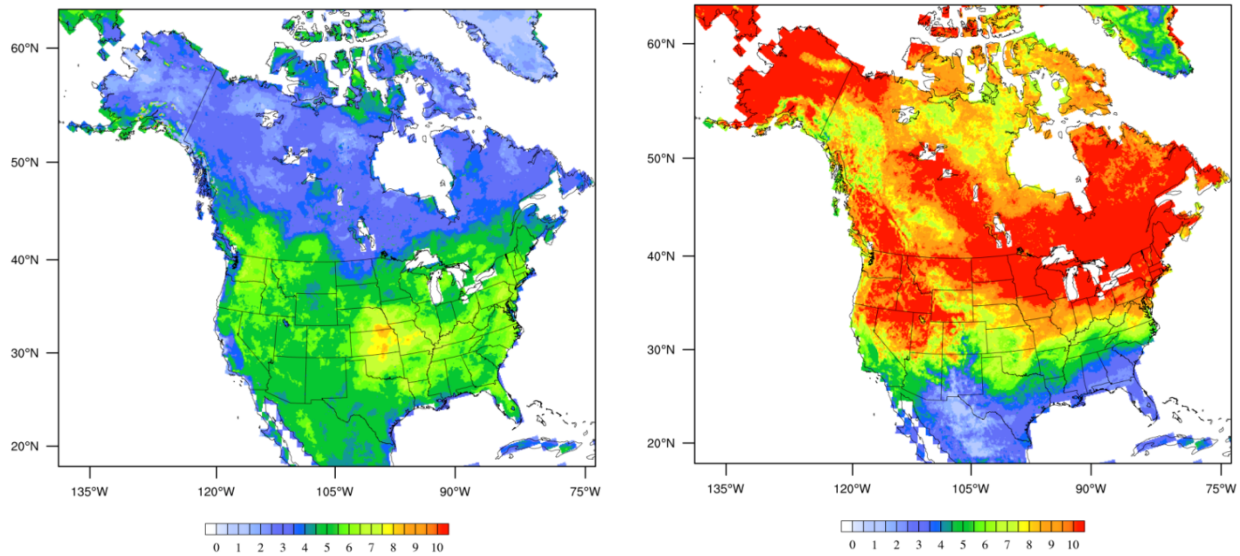

FIGURE 26 Projected Changes in Annual Maximum Temperature (left) and Annual Minimum Temperature (right) for the Decade 2085–2094 (for RCP 8.5 scenario) Compared to 1995–2004

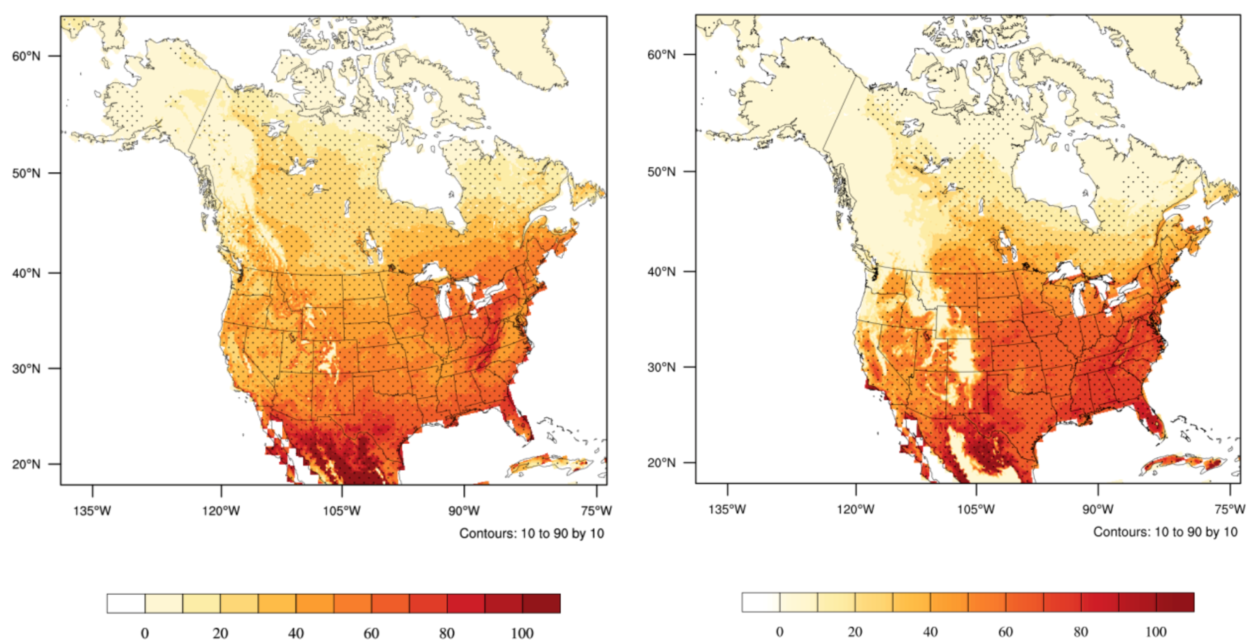

FIGURE 27 Projected Changes in Number of Days (per year) of Tropical Nights (left panel; daily minimum temperature greater than 20°C [68°F]) and Summer Days (right panel; daily maximum temperature greater than 25°C [77°F]) for the Decade 2085–2094 compared to 1995–2004 for the RCP 8.5 Scenario (The dots indicate statistically significant changes based on t-tests.)

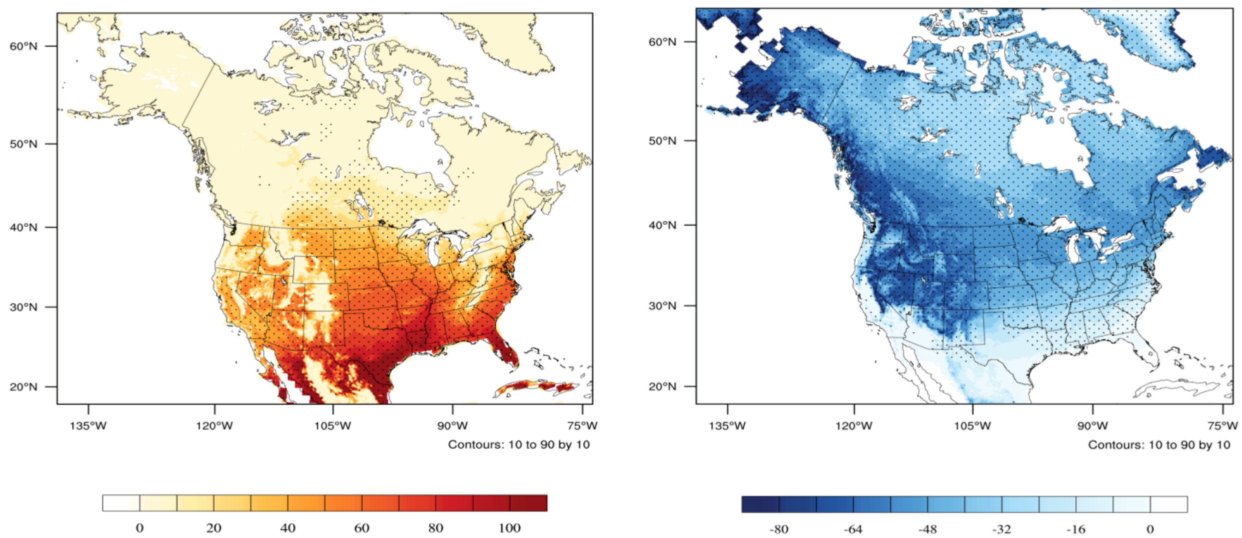

FIGURE 28 Projected Annual Increase in Number of Frost Days Per Year (left; daily minimum temperature less than 0°C) and Hot Days Per Year (right; daily maximum temperature greater than 90°F) for the Decade 2085–2094 (for the RCP 8.5 scenario) Compared to 1995–2004

The left panel of Figure 28 shows the change in the number of days with frost under the RCP 8.5 at the end of the 21st century compared to 1995–2004. Decreases on the order of 80 days are obtained over the western mountainous regions and nearly 30 days over the Midwest. The right panel shows the increase in the number of days with a daily maximum temperature above 90°F. Increases on the order of 100 days occur in the Southwest and increases on the order of 60 days occur over the Midwest.

6 STATISTICAL DOWNSCALING

To enable a comparison of downscaling models, average temperature output at regular 3-hour or 6-hour intervals from a regional climate model must be translated into daily maximum and minimum temperatures (Tmax and Tmin hereafter), recorded by weather stations, and generated by statistical downscaling models. This is not a trivial task, as the amplitude of the daily cycle depends on local factors such as humidity and topography.

The problem in this first task can be generalized as a bandwidth-limited signal that has been sampled at discrete intervals. This signal can be represented completely by taking its discrete Fourier transform (DFT) and retaining all terms up to the Nyquist sampling frequency, which is defined as twice the highest frequency present in the data. The underlying assumption is that the data is bandwidth limited, and sampled at an interval that is at least twice the Nyquist frequency. If these assumptions are valid, then reconstructing the signal via the inverse Fourier transform produces an exact representation of the original signal, not just at the sampled points but at all intermediate points. If, on the other hand, the sequence is sampled below the Nyquist frequency, then the Fourier reconstruction will exactly match the original data at each of the sampled points, but only approximate the signal at intermediate points. The overall RMSE will be proportional to the power of the frequency content above the Nyquist sampling frequency.

We used this approach to estimate temperature profiles at higher sampling rates from 3-hour modeled or sampled temperature profiles. Figure 29 shows a portion of a reconstructed temperature profile. The original data was 5-minute temperature data from a Mesonet site located at Reese Center outside of Lubbock, Texas. It was down-sampled to 3-hour sampling to simulate typical 3-hour observations or climate model output.

The dotted black line in Figure 29 is the original 5-minute data filtered to the Nyquist frequency. The raw data clearly contains higher frequencies, so the resampled signal (red), reconstructed from 3-hour sampling of the original (blue) signal can only approximate the original signal. Note that the reconstructed signal passes through every sampled data point and provides a better approximation of the true signal than simple linear interpolation. The reconstruction was done based on the Fourier transform of an entire year's data, not just the displayed 4 days' data.

On typical days, this approach provides a small improvement in Tmax and Tmin over 3-hour sampled data, as it fills in the likely curve between points. Unless a 3-hour sample point falls exactly at the peak or valley, using Tmax and Tmin from 3-hour data will underestimate Tmax and overestimate Tmin. Tables 5 and 6 show that resampling reduces both the mean and RMSE over 3-hour sampling. The mean error is reduced 60–70% when tested against 5-minute Mesonet observation data. Because the estimation errors are independent of time of day, values are representative of the reconstruction at any time of day, including the peaks and valleys.

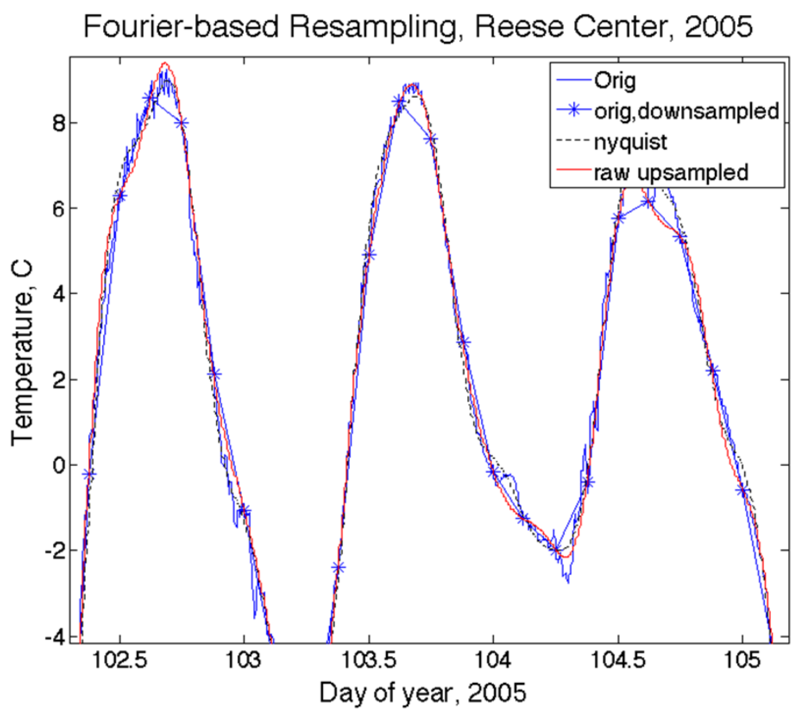

FIGURE 29 Reconstruction of Temperature Profile from 3-hour Sampled Mesonet Temperature Data

TABLE 5 RMSE and Maximum Differences between the Raw 5-minute Data and 3-hour Data for Tmin and Tmax (°C)

	Mean Difference	RMSE	Maximum Difference
Tmax	0.6476	0.8210	3.2300
Tmin	-0.8951	1.2731	5.9240

TABLE 6 RMSE and Maximum Differences between the Raw 5-minute Data and DFT Resampling Data for Tmin and Tmax (°C)

	Mean Difference	RMSE	Maximum Difference
Tmax	0.2816	0.6287	2.8411
Tmin	-0.2848	0.7848	3.7760

Figure 30a (left panels) shows a reconstruction from 3-hour historical data for a typical day in Clovis, New Mexico, for 3-hour WRF data. It shows a small improvement at the peaks and valleys, compared to 3-hour sampling. The top plot shows daily Tmin, Tmax, and daily average temperature using 3-hour data and 5-minute resampled data. The bottom plot shows the reconstructed temperature profile (red) and overlaid on the 3-hour data (blue). A much larger improvement over 3-hour data comes when there is a large temperature swing that occurs at a day-boundary. Figure 30b (right panels) shows that on day 2926, temperature dropped by 15°C between 6 p.m. and midnight. Using 3-hour sampled data results in using the 9 p.m. reading of approximately 268 K, while the true Tmin for the day was actually 10°C lower (~258 K), occurring at midnight. Resampling to 5-minute data picks up a better estimate of the actual daily Tmin, reducing the error to a fraction of a degree C.

The algorithms developed in this task were then applied to the 3-hour regular temperature outputs from WRF simulations to derive the daily maximum and minimum temperatures used in the analysis below.

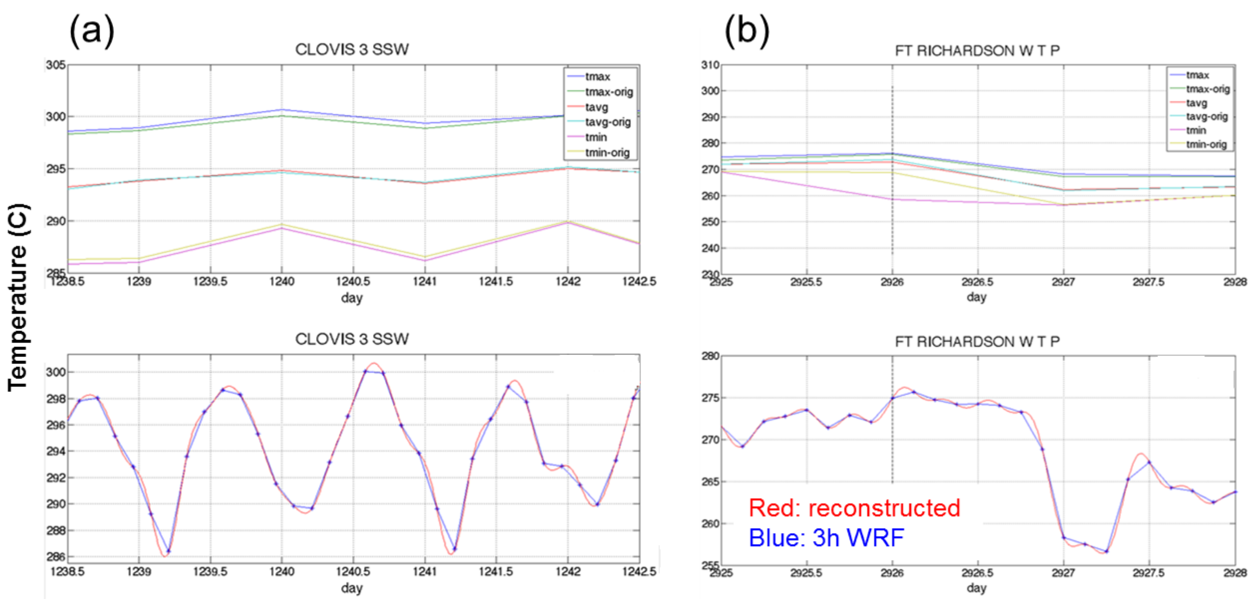

FIGURE 30 Reconstruction of Temperature Profile for (left panel) Clovis, New Mexico, on a Typical Day and (right panel) Fort Richardson, Texas, on a Day with a Large Temperature Swing Occurring at the Boundary of a 24-hour Period

We applied the ARRM model to downscale two different model outputs to the 11 long-term weather stations at DoD installations identified in Figure 2b: the CCSM4 global climate model (Gent et al. 2011) and WRF dynamical downscaling model. The CCSM4 model is the newest version of a long-established and well-documented climate model developed at the National Center for Atmospheric Research. In this study, we used the MOAR run, which saved additional output that is needed to force the WRF model. It has a resolution of 0.9° (latitude) by 1.25° (longitude) and 26 vertical layers in the atmosphere.

6.1 COMPARE OBSERVED, GLOBAL MODEL-SIMULATED, DYNAMICALLY AND STATISTICALLY DOWNSCALED HISTORICAL AND PROJECTED FUTURE TEMPERATURE AND PRECIPITATION

We compare historical simulated and future projected changes in average and extreme temperature and precipitation as generated by five different sources:

1. Observations (OBS),
2. Global climate model simulations (GCM),
3. Global climate model simulations downscaled using the ARRM statistical downscaling model (GCM-SDM),
4. Global climate model simulations downscaled using the WRF dynamical downscaling model (GCM-RCM), and
5. Global climate model simulations downscaled using the WRF dynamical downscaling model and then downscaled again using the ARRM statistical downscaling model (GCM-RCM-SDM).

To compare these sources, we calculated a broad range of climate indicators, including:

- Seasonal and annual mean maximum and minimum temperature and precipitation,
- Daily values above high temperature and precipitation thresholds or below low temperature and precipitation thresholds,
- Wettest day and wettest 5 days of the year, and
- Number of hot/dry and cold/wet days per year.

These indicators were deliberately selected to sample a wide range of seasons and sections of the daily distribution of temperature and precipitation. Although this analysis was

conducted for all 11 weather stations at selected DoD installations in the United States, these results focus on three diverse stations that highlight the diversity of climate conditions across the United States at Watertown, New York; Clovis, New Mexico; and Tacoma, Washington.

This analysis yields a number of interesting, albeit preliminary, results:

- For seasonal mean temperature and precipitation, we find that RCM simulations tend to retain a greater part of the bias of the original GCM forcing than do SDM simulations. When RCM simulations are combined with SDM, this bias is corrected to a large degree (Figure 31). The overall

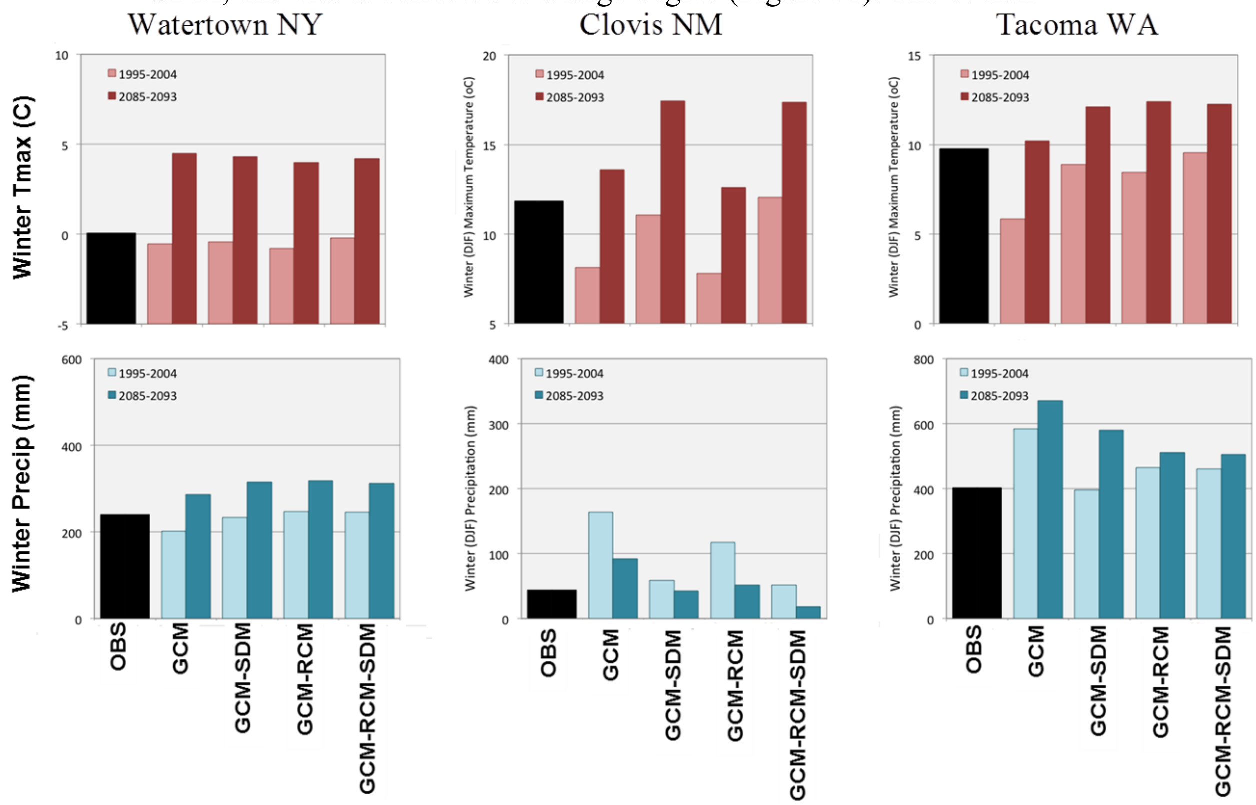

FIGURE 31 Historical and Projected Future Winter (December–February) Maximum Temperature (top) and Precipitation (bottom) at Three DoD Installations Across the United States

- magnitude of change varies slightly but without any consistent difference between the different downscaling approaches.

- For dry days, the GCM consistently underestimates observed values (likely due to the well-known drizzle issue). However, all three types of downscaling correct for this bias and show similar changes in the future (Figure 32).

- For extreme precipitation days (Figure 33), the SDM-only approach agrees with the other approaches for some locations. However, for Tacoma it does

not: here, the SDM-only approach results in projected changes that are more than twice as large as those estimated using other downscaling methods.

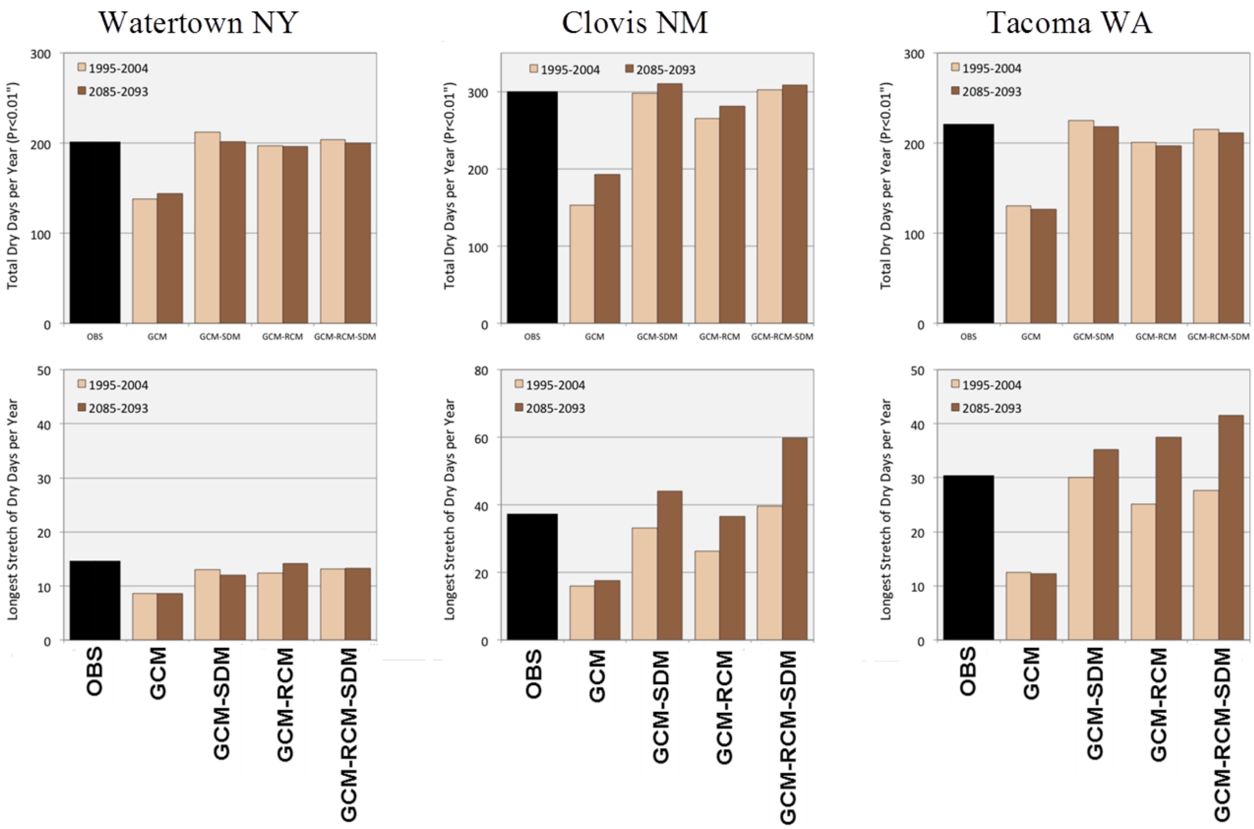

FIGURE 32 Historical and Projected Future Changes in the Total Number of Dry Days per Year (top) and the Longest Period of Dry Days Each Year (bottom)

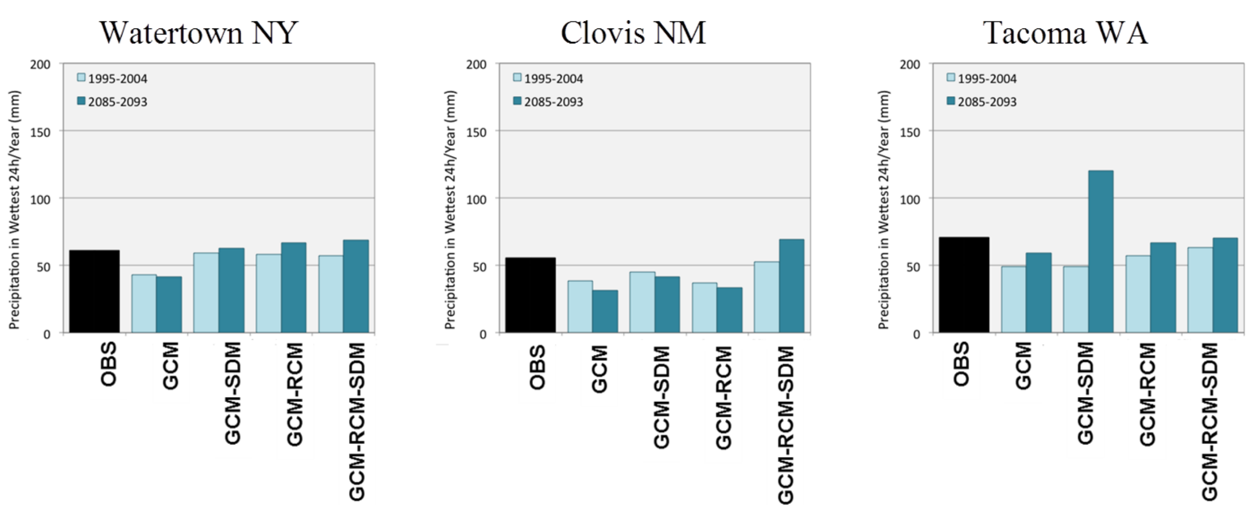

FIGURE 33 Historical and Projected Future Change in Days Per Year with More Than 2 Inches of Precipitation in 24 Hours

7 CONCLUSIONS

We conducted a survey of perceptions regarding climate change and its potential impacts on specific DoD installations. Precipitation changes (particularly at the distribution extremes) at the installations emerged as a primary concern of the participants. The survey results and the priorities identified by the impact assessment community were used to identify a set of climate variables for this downscaling study. Evaluation of different datasets for precipitation and temperature over the model domain resulted in selecting the PRISM dataset as the most accurate data for precipitation and temperature, especially in topographically complex regions. An extensive evaluation of model bias for dynamically downscaled model products during the historical period was generated with observational data over different regions of the CONUS. We establish that the 12-km model resolution and the model setup (parameters, nudging, and spin-up) led to a decrease in model bias as compared to coarser-resolution models, and added value as compared to a method that purely depends on spatial interpolation from a coarser grid. This is especially true when calculating the diurnal variability and extremes of temperature and precipitation.

One of the primary findings of this study is that the accuracy of both relative errors and extreme values is highly dependent on the region being analyzed and the boundary conditions used to drive the simulation. Similarly, knowing the model rank for relative errors from the climatology does not represent how that model performs in extreme climate cases. This could be a result of several different factors. First, the configurations for each climate run are different. While adjusting the initial boundary conditions can be beneficial in many situations, bias correction and nudging are not always an improvement compared to the reference data. In addition, our results show that many variables have the largest errors for surface variables in the wettest and driest regions of the CONUS. High-precipitation regions, such as the Southeast, yield higher errors because of the dominance of convective processes in these regions, which is challenging to predict at this resolution. Similarly, drier regions have been shown to have greater errors or biases due to small-scale processes that are hard to capture using downscaling techniques. The ensemble's ability to capture these historical uncertainties using different reference data is important for future projections across this domain.

The most striking results from the study are that for 2085–2095, the model projections show temperature changes of 5–7°C for summer compared to 1995–2004 and a change of >7°C over northern Canada and Alaska for the winter months. In summer, the projections show a widespread summertime precipitation increase (with precipitation up to 60% higher than present-day average values) throughout much of Canada, Alaska, and the southwestern United States, while the winter experiences lower precipitation than at present over the Southwest and the Southern Great Plains, with precipitation 40–60% lower than present averages. We also estimated the increase in precipitation and temperature in the projections. The model projections indicate 3–5 additional days with precipitation >20 mm/day over the eastern United States, Alaska, and Canada, and ~1 additional day over the western United States. The number of days with precipitation >40 mm/day also increased, especially over the eastern United States and the Cascade Range, with ~2 more days than the present averages. The model projection indicated >60 additional days/year with daily maximum temperature >90°F over the Great Plains and most of the eastern United States (except over the southern mountain ranges). Over the Rockies, the Cascade Range, Alaska, and Canada, <20 additional days had daily maximum temperatures

>90°F. Nearly all of the CONUS and Canada are projected to experience a decrease of >20 frost days/year in 2085–2094 in the RCP8.5 scenario, especially over the West Mountain subregion, which is projected to have >60 fewer frost days/year. The projected changes in extreme temperature show significant elevation dependence, the reasons for which need further investigation.

7.1 IMPLICATIONS FOR FUTURE RESEARCH AND IMPLEMENTATION

High-resolution modeling studies provide stakeholders and the public with knowledge of the uncertainties on a range of climate indicators, including assessing effect on local hydrological processes, surface temperature changes, and heat stress on humans in a warmer climate (Fowler et al. 2007; Buzan et al. 2015). Understanding the strengths and weakness of dynamic downscaling methods is an important step in finding a way to access the risks of future climate and is the primary goal of this research. This type of ensemble downscaling studies can evaluate future uncertainties in societal impacts at spatial scales of interest to the impact assessment and adaptation community (Fowler et al. 2007). We have advanced the state of the knowledge on the use of downscaling products for assessing the impacts of climate change on DoD installations and infrastructure. We have developed a ranking scheme for based on model relative errors from the climatology based on historical observational datasets. Using this ranking matrix, if there is a known overall bias in the dynamically downscaled method for a specific region in all members of the ensemble, that can now be accounted for when making projections of future climate change. Another outcome of the project is that the use of our ensemble could prove valuable in making analyses of uncertainties in projected extreme values. Because most of the uncertainty in future climate comes from choices such as the climate model used and the emission scenario (Déqué et al. 2007), our multi-climate model ensemble, while employing bias correction and spectral nudging, can prove valuable at analyzing the uncertainties in future climate extremes.

7 LITERATURE CITED

Aligo, E.A., W.A. Gallus Jr., and M. Segal, 2007: Summer rainfall forecast spread in an ensemble initialized with different soil moisture analyses. *Weather Forecasting* 22:299–214. doi:10.1175/WAF995.1

Alexander, L.V., Zhang, X., Peterson, T.C., Caesar, J., Gleason, B., Klein Tank, A.M.G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F. and Tagipour, A., 2006. Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research: Atmospheres*, *111*(D5), doi:10.1029/2005JD006290

Alexandru, A., and R. de Elía Laprise, 2007: Internal variability in regional climate downscaling at the seasonal time scale. *Monthly Weather Review* 135:3221–3238

Bowden, J.H., T.L. Otte, C.G. Nolte, and M.J. Otte, 2012: Examining interior grid nudging techniques using two-way nesting in the WRF model for regional climate modeling. *Journal of Climate* 25:2805–2823

Brown, S.J., J. Caesar, and C.A.T. Ferro, 2008: Global changes in extreme daily temperature since 1950. *Journal of Geophyical. Research* 113:D05115. doi:10.1029/2006JD008091.

Bruyère, C.L., Done, J.M., Holland, G.J. and Fredrick, S., 2014. Bias corrections of global models for regional climate simulations of high-impact weather. *Climate Dynamics*, *43*(7-8), pp.1847-1856.

Bukovsky, M.S., and D.J. Karoly, 2007: A brief evaluation of precipitation from the North American regional reanalysis. *Journal of Hydrometeorolog,* 8:837–846.

Bukovsky, M.S., 2011: Masks for the Bukovsky regionalization of North America, Regional Integrated Science Collective, Institute for Mathematics Applied to Geosciences, NCAR, Boulder, CO. Available at http://www.narccap.ucar.edu/contrib/bukovsky.

Buzan, J.R., K. Oleson, and M. Huber, 2015: Implementation and comparison of a suite of heat stress metrics within the Community Land Model, version 4.5. *Geoscientific Model Development* 8:151–170. doi:10.5194/gmd-8-151-2015

Castro, C.L., R.A. Pielke Sr., and J.O. Adegoke, 2007: Investigation of the summer climate of the contiguous U.S. and Mexico using the regional atmospheric modeling system (RAMS). Part I: model climatology (1950–2002). *Journal of Climate* 20:3866–3887.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Monthly Weather Review* 129:569–585. doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2

Chen, M., W. Shi, P. Xie, V.B.S. Silva, V.E. Kousky, R. Wayne Higgins, and J.E. Janowiak, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation. *Journal of Geophysical Research* 113:D04110, doi:10.1029/2007JD009132.

Cheng, W.Y.Y., and W.J. Steenburgh, 2005: Evaluation of surface sensible weather forecasts by the WRF and the Eta models over the western United States. *Weather Forecasting* 20:812–821. doi:10.1175/WAF885.1

Christensen, N.S., A.W. Wood, N. Voisin, D.P. Lettenmaier, and R.N. Palmer, 2004: The effects of climate change on the hydrology and water resources of the Colorado River basin. *Climatic Change* 62:337-363.

Coles, S.G., 2001: An Introduction to Statistical Modeling of Extreme Values, Springer-Verlag, London.

Cook, B.I., and R. Seager, 2013: The response of the North American Monsoon to increased greenhouse gas forcing, *Journal of Geophysical Research* 118:1690–1699. doi:10.1002/jgrd.50111

Craigmile, P.F., and P. Guttorp, 2013: Can a regional climate model reproduce observed extreme temperatures? *STATISTICA* 73(1).

Crane, R.G., and B.C. Hewitson, 1998: Doubled CO_2 precipitation changes for the Susquehanna basin: Down-scaling from the GENESIS general circulation model. *Inernational Journal of Climatology*
18:65–76.

Daly, C., R.P. Neilson, and D.L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Appied. Meteorology* 33:140–158.

Daly, C., G.H. Taylor, and W.P. Gibson, 1997: The PRISM approach to mapping precipitation and temperature. Preprints, 10th Conf. on Applied Climatology, Reno, NV, American Meteorological Society, 10–12.

Daly, C., M. Halbleib, J.I. Smith, W.P. Gibson, M.K. Doggett, G.H. Taylor, J. Curtis, and P.P. Pasteris, 2008: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Internationa. Journal of Climatology* 28:2031–2064.

Déqué, M., Rowell, D.P., Lüthi, D., Giorgi, F., Christensen, J.H., Rockel, B., Jacob, D., Kjellström, E., De Castro, M. and van den Hurk, B.J.J.M., 2007. An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. *Climatic Change*, *81*(1), pp.53-70.

Dettinger M.D., D.R. Cayan, M.K. Meyer, and A.E. Jeton, 2004: Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American river basins, Sierra Nevada, California, 1900–2099. *Climatic Change* 62:283–317.

Done, James M., Greg J. Holland, Cindy L. Bruyère, L. Ruby Leung, and Asuka Suzuki-Parker. "Modeling high-impact weather and climate: lessons from a tropical cyclone perspective." *Climatic Change* 129, no. 3-4 (2015): 381-395.

Donner, L.J., Wyman, B.L., Hemler, R.S., Horowitz, L.W., Ming, Y., Zhao, M., Golaz, J.C., Ginoux, P., Lin, S.J., Schwarzkopf, M.D. and Austin, J., 2011. The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *Journal of Climate*, *24*(13), pp.3484-3519.

Ekström, M., H.J. Fowler, C.G. Kilsby, and P.D. Jones, 2005: New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 2. Future estimates and use in impact studies. *Journal of Hydrology* 300:234–251. doi:10.1016/j.jhydrol.2004.06.017

Fowler, H.J., Blenkinsop, S. and Tebaldi, C., 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International journal of climatology*, *27*(12), pp.1547-1578.

Francis, J.A., and S.J. Vavrus, 2012: Evidence linking Arctic amplification to extreme weather in midlatitudes. *Geophysical Research Letters* 39. doi:10.1029/2012GL051000

Frumhoff, P.C., J.J. McCarthy, J.M. Melillo, S.C. Moser, and D.J. Wuebbles, 2007: Confronting Climate Change in the U.S. Northeast: Science, Impacts, and Solutions. Synthesis report of the Northeast Climate Impacts Assessment (NECIA), Union of Concerned Scientists, Cambridge, MA.

Gao, Y., J.S. Fu, J.B. Drake, Y. Liu, and J.-F. Lamarque, 2012: Projected changes of extreme weather events in the eastern United States based on a high resolution climate modeling system. *Environmental Research Letters.* 7(4):044025. doi:10.1088/1748-9326/7/4/044025.

Gent, P.R., G. Danabasoglu, L.J. Donner, M.M. Holland, E.C. Hunke, S.R. Jayne, D.M. Lawrence, R.B. Neale, P.J. Rasch, M. Vertenstein, P.H. Worley, Z.-L. Yang, and M. Zhang, 2011: The Community Climate System Model version 4. *Journal of Climate* 24:4973–4991. doi:10.1175/2011JCLI4083.1.

Giorgi, F., and L.O. Mearns, 1991: Approaches to regional climate change simulation: A review. *Reviews of Geophysics.* 29:191–216.

Giorgi, F., and L.O. Mearns, 1999: Introduction to special section: Regional climate modeling revisited. *Journal of Geophyical. Research*, 104, 6335–6352.

Giorgi, F., and L.O. Mearns, 2002: Calculation of Average, Uncertainty Range, and Reliability of Regional Climate Changes from AOGCM Simulations via the "Reliability Ensemble

Averaging” (REA) Method. *J. Climate* 15:1141–1158. doi:http://dx.doi.org/10.1175/1520-0442(2002)015<1141:COAURA>2.0.CO;2

Gleckler, P.J., K.E. Taylor, and C. Doutriaux, 2008: Performance metrics for climate models. *Journal of Geophysical Research: Atmospheres* 113. doi:10.1029/2007JD008972

Grell, G.A., L. Schade, R. Knoche, A. Pfeiffer, and J. Egger, 2000: Nonhydrostatic climate simulations of precipitation over complex terrain. *Journal of Geophysical Research* 105(D24):29595–29608.

Grell, G.A., and D. Dévényi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophysical Research Letters* 29. doi:10.1029/2002GL015311

Gutowski Jr, W.J., Arritt, R.W., Kawazoe, S., Flory, D.M., Takle, E.S., Biner, S., Caya, D., Jones, R.G., Laprise, R., Leung, L.R. and Mearns, L.O., 2010. Regional extreme monthly precipitation simulated by NARCCAP RCMs. *Journal of Hydrometeorology*, *11*(6), pp.1373-1379.

Han, J., and J.O. Roads, 2004: U.S. climate sensitivity simulated with the NCEP regional spectral model. *Climate Change* 62:115–154.

Hawkins, E., and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions, *Bulletin of the American Meteorologcial Society,* 90:1095, doi:10.1175/2009BAMS2607.1

Hay L.E., R.L. Wilby, and G.H. Leavesley, 2000: A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States. *Journal of the American Water Resources Association* 36: 387–397. doi:10.1111/j.1752-1688.2000.tb04276.x

Hayhoe K., D. Cayan, C.B. Field, P.C. Frumhoff, E.P. Maurer, N.L. Miller, S.C. Moser, S.H. Schneider, K. Nicholas Cahill, E.E. Cleland, L. Dale, R. Drapek, R.M. Hanemann, L.S. Kalkstein, J. Lenihan, C.K. Lunch, R.P Neilson, S.C. Sheridan, and J.H. Verville, 2004: Emissions pathways, climate change, and impacts on California. *Proceedings of the National Academy of Sciences* 101:12422–12427.

Hayhoe, K., D.J. Wuebbles, and the Climate Science Team, 2008: *Climate Change and Chicago: Projections and Potential Impacts*. Report for the city of Chicago. Available at http://www.chicagoclimateaction.org/pages/research___reports/48.php.

Hayhoe, K., M. Robson, J. Rogula, M. Aufhammer, N. Miller, J. VanDorn, and D. Wuebbles, 2010a: An Integrated framework for quantifying and valuing climate change impacts on urban energy and infrastructure: A Chicago case study. *Journal of Great Lakes Research* 36(sp2):94–105.

Hayhoe K., J. VanDorn, T.C. Croley II, N. Schlegal, and D. Wuebbles. 2010b: Regional climate change projections for Chicago and the US Great Lakes. *Journal of Great Lakes Research* 36:7–21. doi:10.1016/j.jglr.2010.03.012

Hewitson, B.C., and R.G. Crane, 1996: Climate downscaling: techniques and application. *Climate Research* 7:85–95.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones, and A. Jarvis, 2005: Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25:1965–1978. doi:10.1002/joc.1276

Hurrell, J.W., J. J. Hack, D. Shea, J.M. Caron, and J. Rosinski, 2008: A new sea surface temperature and sea ice boundary dataset for the community atmosphere model. *Journal of Climate* 21:5145–5153. doi:10.1175/2008JCLI2292.1

Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shephard, M.W., Clough, S.A. and Collins, W.D., 2008. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research: Atmospheres*, *113*(D13).

Janssen, E., D.J. Wuebbles, K.E. Kunkel, S.C. Olsen, and A. Goodman, 2013: Observational- and model-based trends and projections of extreme precipitation over the contiguous United States. *Earth's Future* 2. doi:10.1002/2013EF000185

Jankov, I., W.A. Gallus, M. Segal, B. Shaw, and S.E. Koch, 2005: The impact of different WRF model physical parameterizations and their interactions on warm season MCS rainfall. *Weather Forecasting* 20:1048–1060. doi:http://dx.doi.org/10.1175/WAF888.1

Jones, C.D., Hughes, J.K., Bellouin, N., Hardiman, S.C., Jones, G.S., Knight, J., Liddicoat, S., O'Connor, F.M., Andres, R.J., Bell, C. and Boo, K.O., 2011. The HadGEM2-ES implementation of CMIP5 centennial simulations. *Geoscientific Model Development*, *4*(3), pp.543-570.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J. and Zhu, Y., 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American meteorological Society*, *77*(3), pp.437-471.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J.J. Hnilo, M. Fiorino, and G.L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bulletin of American Meteorologicl Society* 83:1631–1643. doi:http://dx.doi.org/10.1175/BAMS-83-11-1631

Kattenberg, A., F. Giorgi, H. Grassl, G.A. Meehl, J.F.B. Mitchell, R.J. Stouffer, T. Tokioka, A.J. Weaver, and T.M.L. Wigley, 1996: Chapter 6. Climate Models—Projections of Future Climate, in *Climate Change 1995—The Science of Climate Change*. J.T. Houghton, L.G. MieraFilho, B.A. Chandler, N. Harris, A. Kattenberg, and K. Maskell, Eds., Cambridge University Press, p. 285–358.

Katz, R., P.F. Craigmile, P. Guttorp, M. Haran, B. Sanso, and M. Stein, 2013: Uncertainty analysis in climate change assessments. *Nature Climate Change* 3:769–771.

Klein Tank, A.M.G., F.W. Zwiers, and X. Zhang, 2009: *Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation*, World Meteorological Organization, Climate Data and Monitoring WCDMP-No. 72.

Koenker, R., and G. Bassett Jr., 1978: Regression quantiles. *Econometrica* 46:33–50.

Krishnamurti, T.N., C.M. Kishtawal, D.W. Shin, and C.E. Williford, 2000: Multi-model superensemble forecasts for weather and seasonal climate. *Journal of Climate* 13:4196–4216.

Liang, X.-Z., L. Li, K.E. Kunkel, M. Ting, and J. Wang, 2004a: Regional climate model simulation of U.S. precipitation during 1982–2002. Part I: Annual Cycle. *J. Climate* 17:3510–3529.

Liang, X.-Z., L. Li, K.E. Kunkel, M. Ting, and J.X.L. Wang, 2004b: Regional climate model simulation of U.S. precipitation during 1982–2002, Part 1: Annual cycle. *Journal of Climate* 17:3510–3528.

Liang, X.-Z., J. Pan, J. Zhu, K.E. Kunkel, J.X.L. Wang, and A. Dai, 2006: Regional climate model downscaling of the US. summer climate and future change, *Journal of Geophysical Research* 111:D10108. doi:10.1029/2005JD006685

Liu, C., K. Ikeda, G. Thompson, R. Rasmussen, and J. Dudhia, 2011: High-resolution simulations of wintertime precipitation in the Colorado headwaters region: Sensitivity to physics parameterizations. *Monthly Weather Review* 139:3533–3553. doi:http://dx.doi.org/10.1175/MWR-D-11-00009.1

Liu, P., A.P. Tsimpidi, Y. Hu, B. Stone, A.G. Russell, and A. Nenes, 2012: Differences between downscaling with spectral and grid nudging using WRF. *Atmospheric Chemistry and Physics* 12:3601–3610

Loikith, P.C., B.R. Lintner, J. Kim, H. Lee, J.D. Neelin, D.E. Waliser, 2013: Classifying reanalysis surface temperature probability density functions (PDFs) over North America with cluster analysis. *Geophysial Research Letters* 40:3710–3714. doi:10.1002/grl.50688

Lucas-Picher, P., D. Caya, R. de Elía, and R. Laprise, 2008: Investigation of regional climate models' internal variability with a ten-member ensemble of 10-year simulations over a large domain. *Climate Dynamics* 31:927–940.

Martynov, A., Laprise, R., Sushama, L., Winger, K., Šeparović, L. and Dugas, B., 2013. Reanalysis-driven climate simulation over CORDEX North America domain using the Canadian Regional Climate Model, version 5: model performance evaluation. *Climate Dynamics*, *41*(11-12), pp.2973-3005.

Maurer, E.P., and H.G. Hidalgo, 2008: Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrology and Earth System Sciences* 12:551–563. doi:10.5194/hess-12-551-2008

Maurer, E.P., A.W. Wood, J.C. Adam, D.P. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. *Journal of Climate* 15:3237-3251. doi:http://dx.doi.org/10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2

McGregor, J.L., 1997: Regional climate modelling. *Meteorology and Atmospheric Physics* 63:105–117.

Melillo, J.M., Richmond, T.T. and Yohe, G.W., 2014. Climate change impacts in the United States. *Third National Climate Assessment*.

Mesinger, F., DiMego, G., Kalnay, E. and Mitchell, K., 2006. North American regional reanalysis. *Bulletin of the American Meteorological Society*, *87*(3), p.343.

Montz, B.E., and E. Gruntfest, 2002: Flash flood mitigation: recommendations for research and applications. *Global Environmental Change Part B: Environmental Hazards* 4:15–22. doi:10.1016/S1464-2867(02)00011-6

Morrison, H., Thompson, G. and Tatarskii, V., 2009. Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Monthly Weather Review*, *137*(3), pp.991-1007.

Noh, Y., W.G. Cheon, S.Y. Hong, and S. Raasch, 2003: Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data. *Boundary-layer Meteorology* 107:401–427.

Otte, T.L., C.G. Nolte, M.J. Otte, and J.H. Bowden, 2012: Does nudging squelch the extremes in regional climate modeling? *Journal of Climate* 25:7046–7066.

Palmer, T.N., 2000: Predicting uncertainty in forecasts of weather and climate. *Reports on Progress in Physicsl* 63:71. doi:10.1088/0034-4885/63/2/201

Pan, Z., J.H. Christensen, R.W. Arritt, W.J. Gutowski Jr., E.S. Takle, and F. Otieno, 2001: Evaluation of uncertainties in regional climate change simulations. *Journal of Geophysical Research* 106:17735–17751.

Pan, Z., R.W. Arritt, E.S. Takle, W.J. Gutowski Jr., C.J. Anderson, and M. Segal, 2004: Altered hydrologic feedback in a warming climate introduces a “warming hole.” *Geophys. Res. Lett.* 31:L17109. doi:10.1029/2004GL020528

Ricketts, T.H., 1999. *Terrestrial ecoregions of North America: a conservation assessment* (Vol. 1). Island Press.

Ries, H., and K.H. Schlünzen, 2009: Evaluation of a mesoscale model with different surface parameterizations and vertical resolutions for the Bay of Valencia. *Monthly Weather Rev*iew 137:2646–2661. doi:http://dx.doi.org/10.1175/2009MWR2836.1

Rothfusz, L., 1990: The heat index (or, more than you ever wanted to know about heat index), Technical Attachment SR 90-23. Fort Worth: Scientific Services Division, National Weather Service.

Ruiz, J.J., C. Saulo, and J. Nogués-Paegle, 2010: WRF model sensitivity to choice of parameterization over South America: validation against surface variables. *Monthly Weather Review* 138:3342–3355. doi:http://dx.doi.org/10.1175/2010MWR3358.1

Shepard, D.S., 1984: “Computer mapping: the SYMAP interpolation algorithm” pp. 133–145 in *Spatial Statistics and Models*, G.L. Gaile and C.J. Willmott (eds.), Dordrecht.

Sherwood, S.C., S. Bony, and J.-L. Dufresne, 2014: Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature* 505:37–42. doi:10.1038/nature12829

Smith, R.L., C. Tebaldi, D.W. Nychka, and L.O. Mearns, 2009: Bayesian modeling of uncertainty in ensembles of climate models. *J. Am. Stat. Assoc. Appl. Case. Stud.* 104:97–116.

Steadman, R.G., 1979: The assessment of sultriness. Part I: A temperature-humidity index based on human physiology and clothing science. *Journal of Applied Meteorology* 18:861–873. doi:http://dx.doi.org/10.1175/1520-0450(1979)018<0861:TAOSPI>2.0.CO;2

Stoner, A.M.K., K. Hayhoe, and D.J. Wuebbles, 2009: Assessing general circulation model simulations of atmospheric teleconnection patterns. *Journal of Climate* 22:4348–4372. doi:10.1175/2009JCLI2577.1

Stoner, A.M.K., K. Hayhoe, X. Yang, and D.J. Wuebbles, 2012: An asynchronous regional regression model for statistical downscaling of daily climate variables. *International Journal of Climatol*ogy 33:2473–2494.

Taylor, K.E., 2001: Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres* 106:7183–7192. doi:10.1029/2000JD900719

Torres-Alavez, A., T. Cavazos, and C. Turrent, 2014: Land–sea thermal contrast and intensity of the North American Monsoon under climate change conditions. *Journal of Climate.* 27:4566–4580. doi:10.1175/JCLI-D-13-00557.1

Tebaldi, C., L. Mearns, D. Nychka, and R. Smith, 2004: Regional probabilities of precipitation change: a Bayesian analysis of multimodel simulations. *Geophys. Res. Lett.* 31:L24 213. doi:10.1029/2004GL021276.

Tebaldi, C., R. Smith, D. Nychka, and L. Mearns, 2005: Quantifying uncertainty in projections of regional climate change: a Bayesian approach to the analysis of multi-model ensembles. *Journal of Climate,* 18, 1524–1540. doi:10.1175/JCLI3363.1

Uppala, S.M., Kållberg, P.W., Simmons, A.J., Andrae, U., Bechtold, V.D., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A. and Li, X., 2005. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, *131*(612), pp.2961-3012.

USGCRP (U.S. Global Change Research Program), 2009: Global climate change impacts in the United States, Washington, D.C. Available at http://www.globalchange.gov.

von Storch, H., E. Zorita, and U. Cubasch, 1993: Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. *Journal of Climate* 6:1161–1171.

Walsh, J., D. Wuebbles, K. Hayhoe, K. Kunkel, R. Somerville, G. Stephens, P. Thorne, R. Vose, M. Wehner and J. Willis, 2013: Chapter 2: Our Changing Climate, USGCRP, the Third National Climate Assessment.

Wang, Y., L.R. Leung, J.L. McGregor, D.-K. Lee, W.-C. Wang, Y. Ding, and F. Kimura, 2004: Regional climate modeling: Progress, challenges and prospects. *Journal Meteorological Society Japan* 82:1599–1628.

Wang, J., and V. Kotamarthi, 2013: Assessment of dynamical downscaling in near-surface fields with different spectral nudging approaches using the Nested Regional Climate Model (NRCM). *Journal of Applied Meteorology and Climatology* 52:1576–1591.

Wang, J. and Kotamarthi, V.R., 2014. Downscaling with a nested regional climate model in near-surface fields over the contiguous United States. *Journal of Geophysical Research: Atmospheres*, *119*(14), pp.8778-8797.

Wang, J. and Kotamarthi, V.R., 2015. High-resolution dynamically downscaled projections of precipitation in the mid and late 21st century over North America. *Earth's Future*, *3*(7), pp.268-288.

Wang, J., Han, Y., Stein, M.L., Kotamarthi, V.R. and Huang, W.K., 2016. Evaluation of dynamically downscaled extreme temperature using a spatially-aggregated generalized extreme value (GEV) model. *Climate Dynamics*, pp.1-17.

Wehner, M.F., 2013: Very extreme seasonal precipitation in the NARCCAP ensemble: model performance and projections. *Climate Dynamics* 40:59–80. doi:10.1007/s00382-012-1393-1

Widmann, M., and C.S. Bretherton, 2000: Validation of mesoscale precipitation in the NCEP reanalysis using a new gridcell dataset for the northwestern United States. *Journal of Climate* 13:1936–1950. doi:http://dx.doi.org/10.1175/1520-0442(2000)013<1936:VOMPIT>2.0.CO;2

Wilby, R.L., S.P. Charles, E. Zorita, B. Timbal, P. Whetton, and L.O. Mearns, 2004: *Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods.* The IPCC Data

Distribution Centre, University of East Anglia, UK. Available at http://ipcc-ddc.cru.uea.ac.uk/guidelines.

Wilks, D.S., and R.L. Wilby, 1999: The weather generation game: A review of stochastic weather models. *Progress in. Physical Geography.* 23:329–357.

Wood, A.W., L.R. Leung, V. Sridhar, and D.P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change* 62:189–216. doi:10.1023/B:CLIM.0000013685.99609.9e.

Xie, P., A. Yatagai, M. Chen, T. Hayasaka, Y. Fukushima, C. Liu, and S. Yang, 2007: A gauge-based analysis of daily precipitation over East Asia, *Journal of Hydrometeorology*. 8:607–626.

Xu, Z., and Z.-L. Yang, 2012: An improved dynamical downscaling method with GCM bias corrections and its validation with 30 years of climate simulations. *Journal of Cliamate,* 25:6271–6286.

APPENDIX A

Department of Defense
SERDP-Argonne Project #RC-2242
Dr. Rao Kotamarthi, Lead Principal Investigator

Final Report

Use of Weather and Climate Change Information by DoD Installation Stakeholders

David Changnon and Stephen Strader
Meteorology Program
Department of Geography
Northern Illinois University
DeKalb, IL 60115

June 2013

CONTENTS

Use of Weather and Climate Change Information by DoD Installation Stakeholders

David Changnon and Stephen Strader
Meteorology Program
Department of Geography
Northern Illinois University
DeKalb, IL 60115

ABSTRACT

This study sought to identify the weather and climate change information uses and needs of Department of Defense (DoD) Installation decision makers in the United States. The primary means for obtaining this information came through the dissemination of a questionnaire to DoD stakeholders located at eight Army installations, one Air Force installation, and one Marine installation. This process was facilitated through DoD liaisons working with the SERDP Climate Change projects. Thirty-four questionnaires were completed and returned. Two types of information were considered: use of weather and historical climate data to guide current decisions and use of climate change projections in future endeavors.

Weather directly impacted 33 of 34 stakeholders. It appears that weather information and short-term forecasts are currently being used extensively in daily to weekly decision making at all installations. Uses varied among the different stakeholders (e.g., environmental, sustainability management, conservation, operations and management, emergency managers, and master planning). However, when asked to identify weather extremes that directly impacted their activities and decisions, stakeholders indicated that heavy short-duration rainfall/flooding events and drought/heat waves created the greatest number of installation impacts. When asked whether they used historical climate information to determine how frequently these extremes occur, most said “no” while others generally provided anecdotal information.

Climate change estimates and model projections were not provided to stakeholders or being used in current decision making efforts at most installations (31 of 34 participants). The primary reason for non-use related to the specific “mission” of the stakeholder group. Most (19 of 34) indicated that if climate change estimates were made available they would not incorporate that information into current or future decisions. For those who could see potential use of such projections, most wanted future precipitation projections (i.e., creating conditions that would be too wet or too dry). Most stakeholders were not comfortable addressing issues related to the accuracy of model projections or how to deal with the uncertainty that comes with probabilistic information. Hindrances to use such as scientific uncertainty, the lack of integrative models (e.g., hydrologic, fire, etc.) that use climate change estimates in risk analysis decision processes, and the lack of support from others at the installation were noted by a few stakeholders.

Those involved with this aspect of the project wish they had been able to meet face-to-face with participants to discuss the questionnaire. It was evident by answers provided that a number of stakeholders did not clearly understand the differences between weather forecasts and long-term climate change model projections. Higher level decision makers should have been involved in these discussions. Future assessments of DoD stakeholders need to budget more time and resources to enhance the exchange of knowledge between scientists and users.

Chapter 1. Introduction

Major weather events and a highly variable climate over the past 20 years have created large economic impacts for many weather-sensitive decision makers in the United States and around the world (Changnon and Changnon, 1999; Changnon, 2008). Layered on top of these current issues are concerns related to climate change projections associated with increased levels of atmospheric greenhouse gases (Kunkel et al., 2013). Understanding how weather-sensitive stakeholders located at Department of Defense (DoD) installations deal with current weather issues as well as use climate change projections was the primary goal of this aspect of the SERDP-Argonne project.

Over the past 30 years there has been growing interest in learning how weather-sensitive individuals, organizations, and institutions use available weather and climate data and information to improve decisions (NRC, 1981; Changnon and Fosse, 1981; Changnon and Vonnahme, 1986; Sonka et al., 1992; Changnon, 1992; Changnon et al., 1995, Changnon, 2004). Furthermore, recent studies have examined ways to improve the movement of information between climate scientists and decision makers (Changnon et al, 1984; Pielke, 1997; Morss et al., 2005). As decision makers evaluate ways to minimize or manage weather-related risks they face in their operational or long-term planning decisions, they need to become better aware of the large number of weather and climate resources available to them.

As climate information, predictions, and services have improved (DeGaetano et al., 2010) so have the models and decision support tools that have been developed in collaborative efforts between atmospheric scientists and private-/public-sector users (Dutton, 2002; Changnon and Changnon, 2010). This has led to a more seamless process where decisions that involve some level of uncertainty incorporate probabilistic information into decision tools and models (Morss et al., 2005; NRC, 2006).

This initial stage in the SERDP-Argonne project is to develop a greater understanding of how DoD installation stakeholders are impacted by weather and whether they incorporate state-of-the-art climate change model projections into their decision making efforts. This information gathering will be conducted through the use of a questionnaire. Due to limited resources (e.g, available time for stakeholders, etc.) the stakeholders will either participate in a phone dialogue regarding the questionnaire or will complete the questionnaire on their own time and return it via the internet. The information gained from this dialogue between atmospheric scientists and DoD stakeholders will be useful to the climate modelers involved in the SERDP-Argonne project as they determine what weather variables they want the regional climate models to estimate.

This final report describes the outcome from this information gathering process. It will highlight weather and climate change issues important to DoD stakeholders. We will provide a series of "lessons learned" from this experience at the end of the report.

Chapter 2. The Present Study: Investigation of Weather and Climate Change Uses by DoD Installation Stakeholders

Background

This three-year DoD pilot project (i.e., SERDP-Argonne RC-2242) was funded to examine issues related to future climate change scenarios and implications of those scenarios on the operation, maintenance, and long-term planning decisions at the installation level. The project is divided into various tasks; however, this initial part of the project is essential for climate scientists to understand how a range of DoD installations are currently impacted by weather and climate extremes and how that list of weather-related issues may change in a future climate scenario.

Objectives of the Study

The goal of this initial aspect of the project is to ascertain how various weather extremes (e.g., heavy precipitation events, hot spells, very high/low humidity levels, storms, etc.) impact decisions that stakeholders at selected DoD installations make on a temporal scale that ranges from minutes to years (e.g., evacuating the installation during a record rainfall to increasing the height of dams/levees that protect you from expected increased flooding events). Global climate models now have the capability to determine, with some uncertainty, how the magnitude and frequency of certain types of weather extremes will change in the future at regional scales. This information can then be examined for DoD installations located across the U.S. Those who generate regional climate model output are seeking information from those stakeholders who are directly impacted by a fluctuating/changing climate. The findings should be of value to DoD stakeholders by revealing new opportunities to use and exploit existing model output information more effectively. Further, the decision-makers at the selected DoD installations have the opportunity for direct input into the development of specific modeled data that might be more useful to various applications at their installations in the future.

The information about key weather variables and potential usage of forecasted climate information into decision making efforts will be gathered through in-depth phone interviews/email questionnaires, conducted with the DoD stakeholders (i.e., participants) for each selected DoD installation. The primary tasks associated with this aspect of the project will:

1. Assess current weather extreme issues (e.g., heavy rain events, storms, hot/humid weather, etc.) that currently impact a range of decisions at each DoD installation.
2. Assess the relative value assigned to the use of future climate model output data and information (given the current level of uncertainty with regional climate model output).
3. Ascertain whether DoD decision makers are able to obtain critical information regarding these events and the frequency of their occurrence (now and using future climate scenarios).
4. Determine impediments to using existing data/information about future weather extremes.

5. Determine what types of climate/weather data/information they most understand/trust and how they want it communicated to them.

After the data has been gathered from all the DoD installations sampled, it will be analyzed for current usage, problems with current modeled output data, and future desires/needs as it applies to weather-related decisions. Information will be compared among those installations sampled to identify similarities and differences. The identity of the participants will remain anonymous throughout the project.

Questionnaire Development

The primary goal of the questionnaire (See Appendix A) was to develop a greater understanding of how weather-sensitive decision makers are impacted by weather and whether they have incorporated climate change projections into their decision efforts. Similar to previous questionnaires/surveys developed by the investigators (Changnon et al., 1995; Changnon, 2004), the questionnaire was laid out in five parts. The questionnaire was developed so that whether it was completed via a phone interview or by the participant (in isolation) and then emailed back, it would take less than an hour of participant's time.

The first part provides a summary of the project and gives the participant some background on those who developed the questionnaire and are seeking information. By continuing into the questionnaire, the participant is giving his or her informed consent for investigators to use the provided information. Importantly, those leading this aspect of the SERDP-Argonne project expected that the selected participants would have good working definitions for "weather" and "climate." Because nearly all of the questionnaires were completed by a participant on their own time and then returned by email to us, we were concerned about their meteorological background and the ability to differentiate between weather and climate issues.

The second part of the questionnaire focused on learning about the participant, his/her role at the installation, and how their installation responsibilities were affected by weather. We wanted to get some idea about whether they felt that these responsibilities were currently changing or could change in the future, especially if the climate were to change. To help engage them in this aspect of the questionnaire we identified a number of weather extremes with the expectation that one or more of these would get them thinking about specific situations where weather had a significant impact on what they did. Once these weather extremes were noted, we wanted to determine whether they used historical climate data and information to determine the frequency of such extremes, where this historical data was obtained, and whether this data was currently used to make installation decisions.

Part three of the questionnaire wanted to know whether the participants used climate change data, projections, estimates, and other related information generated from climate models. We wanted to know who provided this information to them, how it was applied within their job responsibilities, and what specific data was being used (i.e., projected climatic means or extremes). We hoped they could provide some idea of the "value" associated with the use of climate model output in their decisions. Finally, we wanted to the participant to identify factors that limit the use of climate change information.

The next section of the questionnaire was developed primarily for those who indicated that they did not currently use climate model output in their job responsibilities. We wanted to

think about how they might use climate model output data, and if so, what types of data would be most important to them (e.g., temperature, precipitation, storms, etc.). We also wanted to ascertain the levels of accuracy and uncertainty that they were comfortable working with when using climate model projections. Could more detailed explanations for specific weather anomalies or analogs to similar past events be useful? Finally, we explored the perceived hindrances that would prevent them from ever using climate change estimates in their decisions.

The last section gave participants an opportunity to comment on the questionnaire, speak to various questions they had with this process, go back and revisit sections or seek further information. We wanted them to know that we appreciated them taking time to complete the questionnaire.

Once a draft of the questionnaire was completed in September 2012 it was reviewed by Northern Illinois University's (NIU) Institutional Review Board (IRB). This process was mandatory as the questionnaire would involve human subjects. Changnon received word on October 2, 2012, that the final draft of the questionnaire was approved for use and dissemination.

Efforts to Engage DoD Installation Stakeholders

In late October 2012 Changnon and the project's lead P.I., Dr. Rao Kotamarthi (Argonne National Laboratory), began to have discussions with Dr. John Hall (Director, SERDP) about identifying a key liaison for each wing of the Armed Forces who would lay the groundwork for interactions between Changnon and DoD Installation stakeholders. A copy of the questionnaire was provided to each of the selected liaisons involved in the SERDP projects. The liaisons evaluated and approved the questionnaire for use within their wing of the Armed Forces.

In January 2013 these liaisons began to reach out to the military leaders at those installations identified in the SERDP-Argonne project (see Appendix B). This process was long and arduous for all involved and Changnon is grateful for the efforts of all involved. During March 2013 the questionnaire was sent via email to various installation leaders who then disseminated it to those who they identified as being weather-sensitive stakeholders at their installations. All but one stakeholder completed the questionnaire by themselves and returned it to Changnon via email. One participant wanted to discuss the questionnaire over a phone call. The 34 questionnaires from the 10 installations were completed and received by May 3, 2013. Although a larger sample with a greater number of participants from the various weather-sensitive units was desired, given the level of effort required to obtain the 34 completed questionnaires, the researchers were overjoyed by what they had to evaluate. Those who completed the questionnaire fell into one of the following installation activities:

- airfield operations,
- sustainable management/conservation efforts (e.g., game warden, cultural sites, range planning),
- operations and management,
- emergency management,
- environmental issues (e.g., water and air quality compliance, waste water issues, etc.),
- engineering,
- master planning, and
- plans, analysis and integration.

Chapter 3. Results of Questionnaire Survey

Impact of Weather and related extremes on Stakeholders

All but one of the 34 participants who completed the questionnaire indicated that their installation unit was directly impacted by weather in one or more ways. Participants provided very detailed answers in which they described specific activities and/or decisions which involved weather. How participants were impacted by weather was directly related to their unit "missions". For example, emergency managers focused on weather (e.g., hurricanes, flooding, fires, etc.) hazard preparation, response, and recovery, whereas those in environmental units monitored weather conditions, especially those that could impact local habitats or compliance efforts at the installation or nearby. Clearly, these weather-sensitive decision makers were keenly aware that weather on the scale of hours to days into the future could impact their activities and responsibilities on the installation.

Half of the participants indicated that weather-related responsibilities could change over time, however when asked if they were currently changing, nearly half said "no" and 11 of the 34 were not sure or perhaps did not fully understand the question. One example of how responsibilities were changing focused on installation responses to regional weather extremes (i.e., hurricanes). Not only are emergency management units involved in installation recovery efforts, they have become increasingly involved in community and regional recovery efforts (e.g., Hurricane Sandy in October 2012).

When asked to identify those recent weather extremes that most affected their installations (i.e. created numerous impacts for their units), participants generally identified those associated with the climate region where their installation was located (i.e., differences between the climate in the U.S. Southwest versus the U.S. Southeast). Twenty-two participants indicated that heavy short-term duration rainfall events created impacts (i.e., flooding, erosion, habitat issues, etc.). Long-term drought and related heat waves created impacts for 18 participants (i.e., wildfire issues, HVAC problems, training, etc.). Long durations of wet/dry periods or hot periods were identified by 13 participants as a major weather concern (i.e., forest fires, habitats, etc.). Severe weather events such as hurricanes or tornadoes were listed as important by 10 of the 34 participants. Finally, six identified severe winter conditions as impacting their installations (this small number may be related to the location of the selected installations).

Climate background related to Weather Extremes

When participants were asked about their knowledge of weather event frequency, 21 of 34 indicated that they had little or no background on how frequently a major rainstorm or long duration hot period occurred at their installation. A couple participants commented that because they move around within units at the installation or from installation to installation, having time to develop this knowledge didn't exist or wasn't perceived as important to their mission. Some of the answers suggested that participants were having a difficult time separating a weather event from a climatic frequency. This "weather forecast" versus "historical climate data" concern was further apparent when participants were asked where they obtained information about the historical frequency of various weather events. Several participants indicated that they received this information from a weather forecast provided by those on the installation. Those few that

did complete an analysis of historical climate data to determine storm event frequency obtained their data from local installation sources (i.e., available data or publication) or some government agency (not always named). Most (18 of 34) participants indicated that historical climate information was not used in installation decisions. Those who said "yes" frequently demonstrated their confusion between weather forecasts and historical climate data by giving examples of using weather forecast information to make what most would consider a climate-related decision.

Current use of Climate Change information generated from Climate Models

As the questionnaire went from ascertaining how weather impacted the unit's activities and responsibilities to determining how available climate change information was being integrated into current and future decisions it became very clear to those who designed the questionnaire that these participants were far removed from using climate change model output in their weather-related decisions. Not surprising, only 3 of 34 participants indicated that they considered climate change information as they evaluated future weather-related issues that could impact their mission. Those considering use of climate change information indicated that they developed a background in this subject matter based on personal research (i.e., beyond their mission responsibilities). These individuals were generally involved in environmental efforts (i.e., concerned about changing habitats and impacts on various species) and were interested in how projected precipitation and temperature levels would change in the future and thus impact environmental efforts at the installation. Although these few users were concerned about the accuracy of the model projections, the credibility of the sources, and the lack of a "critical mass" of installation staff to work on climate change issues, the primary reason why they did not directly integrate climate change model projections into their decisions is that it was not part of their units "mission" (i.e., to consider uncertain issues that could change in the future). This was the first time, but not the last, that it became apparent to those who developed the questionnaire that we should have talked with higher level decision makers, those who make decisions about unit missions and would highlight the importance to these participants of incorporating climate change model projections into current and future unit decisions.

If Climate Model output is not used, what types of information would Stakeholders like if it were available?

Those who developed the questionnaire were not too surprised by the lack of use of climate model projections, however, we expected to receive more feedback on questions related to the types of climate change information that participants would like to have if it were available to them. Not only did participants not include climate change output in current decisions, most did not have a good idea about how climate change information could be incorporated into their mission activities and responsibilities. When asked how these participants would use climate change information if it were available to them, 15 of 34 said they would consider its use, however, in evaluating their answers in greater depth some of these 15 were actually talking about using weather forecasts not climate projections. When asked what information derived from regional climate models would be useful to them, precipitation data ranked highest (17 of 34), with temperature data second (11 of 34 participants). These results paralleled those from the initial questions focusing on weather-related impacts. Interesting heating degree day (HDD) and

cooling degree day (CDD) information was deemed important by seven participants. Wind speed, humidity levels, and seasonal and long-term drought were identified by four participants as important weather variables.

When asked about the level of accuracy that the participants could live with if they were to use climate change information in their decisions, the vast majority (25 of 34) did not have an answer. Furthermore, 23 of 34 participants did not know how they wanted model-related uncertainty expressed. When asked whether the participants were comfortable using probabilistic information in decisions, 15 of 34 said yes. Answers to these questions brought up another flaw in this limited dialogue between scientist and stakeholder, that is, many participants either did not understand these questions or did not know how to address them as they relate to climate change model output.

Participants were asked if having analogs to similar past events/periods would be informative. Only seven of 34 indicated that analogs would be helpful as they plan for the future. However, in the earlier part of the questionnaire when participants were asked to discuss various weather events that impacted their current decisions, more than seven participants identified a specific flood, drought year(s), or hurricane that created huge impacts for them suggesting that the participants were not comfortable with the word “analog.” When asked if they wanted further explanations for the uncertainty associated with climate change projections only five of 34 thought that would be beneficial information.

What hinders the use of Climate Change Estimates in their Decisions?

When study participants were asked what issues would hinder the use of climate change projections in their decisions, most did not answer. Part of the reason for lack of participation at this point might be related to the fact that most participants had lost interest in answering the questionnaire. Less than 25% of the participants singled out scientific uncertainties, the lack of models to integrate climate change information into decisions, or lack of support from others above them as reasons for not applying climate change estimates in their decisions. Many comments unrelated to this specific question were provided by participants at this point in the questionnaire. One participant noted that marines prepare themselves for any current or future situation and thus that person could not see value in integrating climate change information into their activities. Several noted that they would never use climate change projections. Others commented on fact that incorporating climate change projections were not part of their unit’s mission.

Chapter 4. Conclusions, Implications, and Recommendations

Summary of Motivation and Scope of the Project

The project's goal was to develop an enhanced understanding of how DoD installation stakeholders are impacted by weather and whether they use (and in what ways) climate change model projections. This information would then be informative to atmospheric scientists involved in the development of regional climate model projection output for those at the study's installations. This aspect of the DoD SERDP-Argonne study relied on efforts of many individuals. Like a track relay where the baton is handed from racer to racer, the questionnaire went from those who developed it to DoD liaisons, then it was delivered to the installation leaders who then disseminated it to those who he/she thought might be weather-sensitive and interested in completing it. Similar to a relay, each exchange of questionnaire could be associated with problems. This process represented a good demonstration project (i.e., case study) for those interested in gathering information from weather-sensitive decision makers in the DoD.

Summary and Implications of Present Use of Weather Information

The participants who completed and returned the questionnaire are weather-sensitive. Not only are they impacted by various types of extreme weather conditions, unit decisions that they are involved with are often impacted by weather events. The top weather problems facing participants, no matter what unit they worked in, involved precipitation. Most concerns noted by participants were related to high-intensity, short-duration heavy rainfall events that frequently led to flooding problems on or near the installation. The second most important weather concern mentioned by participants related to long-duration, hot and often dry weather events, followed by extreme and damaging storms (e.g., tropical storms and severe thunderstorm-related issues). These decision makers relied on good weather forecast information to plan, where possible, respond, and recover from these events. Weather forecasts provided by base personnel is reliable and useful in activities faced by these participants as part of their unit's day-to-day mission.

Summary and Implications of Present Use of Climate Change Projections

Climate change projections are only considered by a handful of the participants sampled in this study. Through further reflection on this study, and who participated in this information transfer, we should not be surprised by this outcome. Although these individuals are generally aware of climate change, their unit's mission does not require them to consider its use in future decisions. Those decision makers who determine whether a unit incorporates climate change projections in their decisions should have been asked to answer the questions related to current and projected use of climate change model estimates.

Recommendations for Climate Modelers

An expected outcome from this initial part of the SERDP-Argonne project was to inform climate modelers about the specific weather-related needs of installation stakeholders. Although these stakeholders face many weather-related issues as part of their roles at the installation, at this point most are not expected to incorporate climate change projections into their decisions. Despite this lack of interest in climate change model projections the participants provided some ideas as to what types of weather currently create problems for them at their installation. Although the 10 installations sampled are located across the United States in many different regional climates, a couple themes came out of the completed questionnaires that could be useful to regional climate modelers.

First, high-intensity, short-duration rainfall events created a wide variety of problems for many different types of decision makers at these installations. Climate modelers should examine how projections related to frequency and magnitude of extreme precipitation events change in the future. Second, warm season heat waves (and related droughts) that last several days or much longer impact participants involved in environmental, sustainability, and engineering decisions. Changes to future summer temperature conditions (e.g., duration or frequency of daily high temperatures above a certain threshold) might impact decisions faced by many installation decision makers including those impacted by fires on or near the installations. Finally, meteorological characteristics (wind speed, storm surge, extreme rainfall rates, etc.) associated with extreme storms such as hurricanes and thunderstorms appear to be important to certain decision makers (i.e., emergency management, those who oversee the installation infrastructure, etc.). Importantly, any weather type that curtails or stops troop training efforts is viewed as a major issue.

Conclusions

In most studies that involve some type of questionnaire or survey, outcomes provide insight into the tool used to obtain information, the participants, and the ability to address the initial study goals. Below is a list of "lessons learned" from this study.

- *Although most participants were "weather savvy", climate change was not on their "radar" (i.e., not part of their unit's mission).* The mission for most units as it related to weather was focused at daily operational decisions (i.e., prepare, respond, and recover). Further, because climate change efforts were not part of the participant's day-to-day responsibilities, their direct input into the development of specific modeled data that might be more useful to various applications at their installations was limited. Until the use of climate change information is "valued" by installation leaders and becomes part of stakeholder responsibilities, it won't be considered important.

- *Essential resources (time and money) need to be set aside to conduct face-to-face interviews with DoD installation stakeholders.* Many misconceptions related to this questionnaire could have been resolved through a dialogue between scientist and participant. This DoD study did not contain many "directly" useful climate change results and insights. This could be improved through dialogues with stakeholders.

- *Educating participants on climate change issues (e.g., the science, global versus regional climate models projections, uncertainty in model output, how they may be impacted by changes in climate, etc.) is essential if they and their leaders are to "buy into" the use of model projections in their decisions.* Educate stakeholders, who in this study, frequently confused terms such as "weather" and "climate." Furthermore, most stakeholders are not aware of existing historical climate data bases or where to obtain climate information that could enhance their decisions. If they see the "value" associated with climate data, information and projections, they may be more comfortable using it (i.e., if it improves decisions and saves money/time…good!).

- *Precipitation extremes appear to be the most important weather variable impacting operations at these installations.* Future climate model output that examines the frequency and magnitude of future short-duration, high-intensity rainfalls should be of great interest to a number of installation stakeholders. Heat waves and fire hazards are also important. Most stakeholders do not have decision support models (e.g., hydrologic, etc.) that incorporate weather data into them.

- *Those involved in environmental issues (e.g., flora, habitats, etc.) at all installations sampled may be the best group to test climate model projections with as they must think about future issues.* Working hands-on with one user group in "case study" situation might help create a process to integrate climate change information into decisions that then can be tested and used by other installation stakeholders. Results of this questionnaire indicate that most stakeholders function (based on their related missions) "reactively" to weather situations. For them to incorporate climate change information into future decisions the participants will have to become more "proactive" in their actions.

ACKNOWLEDGMENTS

This aspect of the SERDP-Argonne project originated in February 2011 when Dr. Rao Kotamarthi contacted Changnon to discuss ways to obtain weather use information from DoD stakeholders. Through these discussions a proposal was developed and submitted. We appreciate the guidance and support provided by Dr. John Hall (former RC Program Manager, SERDP) throughout this initial part of the Argonne project (RC-2242).

We also want to acknowledge the assistance provided by the three DoD liaisons who helped establish collaborations with the various installations associated with the Argonne project, Wanda Johnsen (Army), Daniel Kowalczyk (Air Force), and Elmer Ransom (Marine Corps). Without their assistance the ability to obtain any necessary information from DoD installation stakeholders would have been greatly reduced. They provided important feedback on the questionnaire and information related to communicating with stakeholders.

Finally, I want to thank administrators on the NIU campus, especially those in Grants Fiscal, who helped manage this project and its various components.

LITERATURE CITED

Changnon, D., 2004: Improving outreach in atmospheric sciences: Assessment of users of climate products. *Bull. Amer. Meteor. Soc.*, *85*, 601-606.

Changnon, D., and S.A. Changnon, 2010: Major growth in some business related uses of climate information. *J. Applied Meteor. and Climat., 49*, 325-331.

Changnon, S.A, 1992: Contents of climate predictions desired by agricultural decision makers. *J. Appl. Meteor., 31*, 1488-1491.

Changnon, S.A., 2008: Temporal changes in extremely damaging storms. *Phys. Geogr., 29*, 1-10.

Changnon, S.A., and E.R. Fosse, 1981: Impacts and use of climatological information in the hail insurance industry. *In* Climate and risk: Proceedings of a conference sponsored by the Center for Advanced Engineering Study at MIT (Pocinki et al., editors), The MITRE Corporation, McLean, VA, Vol. I: 4.50-4.96.

Changnon, S.A., J.L. Vogel, and W.M. Wendland, 1984: New climate delivery system developed. *EOS, 65,* 326.

Changnon, S.A., and D. Vonnahme, 1986: Use of climate predictions to decide a water management problem. *Water Resour. Bull., 22,* 649-652.

Changnon, S.A., J.M. Changnon, and D. Changnon, 1995: Uses and applications of climate forecasts for power utilities. *Bull. Amer. Meteor. Soc., 76,* 711-720.

Changnon, S.A., and D. Changnon, 1999: Record high losses from weather disasters in the United States during the 1990s: How excessive and why? *Natural Hazards, 13*, 287-300.

DeGaetano, A.T., T.J. Brown, S.D. Hilberg, K. Redmond, K. Robbins, P. Robinson, M. Shulski, and M. McGuirk, 2010: Toward regional climate services. *Bull. Amer. Meteor. Soc., 91,* 1633-1644.

Dutton, J.A., 2002: Opportunities and priorities in a new era for weather and climate services. *Bull. Amer. Meteor. Soc., 83,* 1303-1311.

Kunkel, K.E., T.R. Karl, H. Brooks, J. Kossin, J.H. Lawrimore, D. Arndt, L. Bosart, D. Changnon, S.L. Cutter, N. Doesken, K. Emanuel, P.Y. Groisman, R.W. Katz, T. Knutson, J. O'Brien, C.J. Paciorek, T.C. Peterson, K. Redmond, D. Robinson, J. Trapp, R. Vose, S. Weaver, M. Wehner, K. Walter, and D. Wuebbles, 2013: Monitoring and understanding trends in extreme storms: State of knowledge. *Bull. Amer. Meteor. Soc., 94,* 499-514.

Morss, R.E., O.V. Wilhelmi, M.W. Downton, and E. Gruntfest, 2005: Flood risk, uncertainty, and scientific information for decision making. *Bull. Amer. Meteor. Soc., 86,* 1593-1601.

National Research Council, 1981: Managing climatic resources and risks. National Academy Press, Washington, D.C., 51 pp.

National Research Council, 2006: Completing the forecast: Characterizing and communicating uncertainty for better decisions using weather and climate forecasts. National Academy Press, Washington, D.C., 112 pp.

Pielke, R.A., Jr., 1997: Asking the right questions: Atmospheric science and societal needs. *Bull. Amer. Meteor. Soc., 78*, 255-264.

Sonka, S.T., S.A. Changnon, and S. Hofing, 1992: How agribusiness uses climate predictions: Implications for climate research and provision of predictions. *Bull. Amer. Meteor. Soc., 73,* 1999-2008.

Appendix 1. Questionnaire Survey

Assessment of the Climate-Related Needs of U.S. DoD Decision Makers as it Pertains to Future Regional/Local Climate Scenarios

David Changnon
Board of Trustees Professor
Meteorology Program
Department of Geography
Northern Illinois University

This three-year Department of Defense (DoD) pilot project is being funded to examine issues related to future climate change scenarios and implications of those scenarios on the operation, maintenance, and long-term planning decisions at the installation level. The project is divided into various tasks; however, your participation in the initial part of the project is essential for climate scientists to understand how your installation is currently impacted by climate extremes and how that list of weather-related issues may change in a future climate scenario.

The goal of this initial aspect of the project, which I will lead, is to ascertain how various weather extremes (e.g., heavy precipitation events, hot spells, very high/low humidity levels, storms, etc.) impact decisions that you and others at selected DoD installations make on a temporal scale that ranges from minutes to years (i.e., evacuating the installation during a record rainfall to increasing the height of dams/levees that protect you from expected increased flooding events). Global climate models now have the capability to determine, with some uncertainty, how the magnitude and frequency of certain types of weather extremes will change in the future at regional scales. This information can then be examined for DoD installations located across the U.S. Those who generate regional climate model output are seeking information from those stakeholders who are directly impacted by a fluctuating/changing climate. The findings should be of value to DoD stakeholders by revealing new opportunities to use and exploit existing model output information more effectively. Further, the decision-makers at the selected DoD installations have the opportunity for direct input into the development of specific modeled data that might be more useful to various applications at their installations in the future.

The information about key weather variables and potential usage of forecasted climate information into decision making efforts will be gathered through in-depth phone/email interviews, lasting about an hour, conducted with the environmental coordinator for each selected DoD installation. The primary tasks associated with this aspect of the project will:

- Assess current weather extreme issues (e.g., heavy rain events, storms, hot/humid weather, etc.) that currently impact a range of decisions at each DoD installation.
- Assess the relative value assigned to the use of future climate model output data and information (given the current level of uncertainty with regional climate model output).
- Ascertain whether DoD decision makers are able to obtain critical information regarding these events and the frequency of their occurrence (now and using future climate scenarios).

- Determine impediments to using existing data/information about future weather extremes.
- Determine what types of climate/weather data/information they most understand/trust and how they want it communicated to them.

After the data has been gathered from all the DoD installations sampled, it will be analyzed for current usage, problems with current modeled output data, and future desires/needs as it applies to weather-related decisions. Information will be compared among those installations sampled to identify similarities and differences. To ensure confidentiality, data and information gathered through the interview process will be stored in a locked file cabinet in Changnon's Northern Illinois University office (100 Davis Hall, Department of Geography).

This part of the project will generate a report summarizing the findings and making recommendations to those who will be developing downscaled climate information for greater applicability and for future research. This report will be distributed to participating DoD installations and to Dr. John Hall, Program Director for SERDP (Strategic Environmental Research and Development Program). Furthermore, results from this research effort will be disseminated in one or more public venues (e.g., conference paper, refereed article, etc). *Your willingness to participate in these interviews means that you give me consent to make available summarized research results. No individual interviewed as part of this research effort will be identified in these external reports or presentations. Only general results will be reported. As such participants may skip questions they prefer not to answer. Also, each participant will be asked to verify that they are 18 years of age or older. Again, participation in this aspect of the research (i.e., the phone interview) is voluntary and you may withdraw at any time without penalty.* If there are any questions regarding your rights as a research participant in these phone interviews please contact:

Ms. Jeannette Gommel, Research Compliance Coordinator
Office of Research Compliance
Division of Research and Graduate Studies
Northern Illinois University
DeKalb, IL 60115
Phone: (815) 753-8588
Email: jgommel@niu.edu

If before or after the phone interview you have a need to contact me, below is my contact information:

Dr. David Changnon, Professor
Meteorology Program
Department of Geography
Northern Illinois University
DeKalb, IL 60115
Phone: (815) 753-6835
Email: dchangnon@niu.edu

Outline of Questions for Interviews with DoD Installation Stakeholders

A. Introductory discussion period

1. Discuss procedures **for obtaining informed consent**:
    - *a. Do they understand and approve of the idea that their answers may be incorporated into external reports and presentations disseminated in a public venue? YES or NO*
    - b. *No individual interviewed as part of this research effort will be identified in these external reports or presentations*.
    - c. *Results for individual questions will be generalized for the sample of participants*.
    - d. *Participants may skip questions they prefer not to answer*.
    - *e. I would like you to verify that you (i.e., the participant in the interview) are 18 years of age or older.*

2. Who I/we are, experience/training, organization.

3. Explain my past involvement in assessing uses of climate data/information in a number of weather-sensitive sectors (ag, transportation, energy, insurance, etc.)

4. **DEFINE** terms...climate vs. weather, climate extremes, predictions and regional climate model output, weather-related risk, uncertainty

5. Describe/review the **PROJECT** (review 2 –page description that they will receive by email before the phone interview begins):
    - a. Goals/objectives of the project.
    - b. Why this information dialogue is critical to both communities (science community and DoD administrators/policy makers).
    - c. Identify potential **BENEFITS** to person/installation/DoD (greater efficiency, better design of adaptation plans, better operations, etc.)
    - d. Describe “how” the data are to be collected (**INTERVIEWS** with those having same type of functions within the installations) and analyzed.

6. Describe the **SCOPE** of the interview…the general agenda to be followed (Q & A), and the talking points to be covered.

B. Tell ME about your area of responsibilities in your installation

1. In what ways does the **weather** affect what:
    - ***a.*** **YOU DO**?
    - ***b.*** Your **unit/section** of the installation must do?

2. Could this **CHANGE**? Has it been **CHANGING**?

3. How did certain **RECENT WEATHER EXTREMES** affect you/your installation (i.e., which created impacts and are you prepared to handle these)?

    a. Summer of 2012 drought/excessive heat wave

    b. Unusually cold/snowy winters

    c. Extremely heavy short-term precipitation events (tropical, convective in excess of 2”/day)

    d. Long durations when you have extreme weather conditions (e.g. a stretch of wet days, hot/cold days)

    e. Severe weather (high winds, hail, tornadoes, hurricanes, etc.)

4. Currently, do you have any idea how **<u>frequently</u>** these types of **weather/climate extremes** occur at your installation (e.g., by year, by season, etc.)?

    a. What is the **source** of that information (i.e., did it come from a NOAA/government document, or was an analysis of local meteorological data conducted by DoD staff, etc.)?

    b. Is that **scientific information** used in installation operation/planning decisions? If yes, could you provide some examples?

C. Do you <u>UTILIZE</u> climate change (future) data/estimates/information generated from <u>CLIMATE MODELS</u> (If “no” go to D)

1. If you get and use predicted/projected climate data:

    a. How is ***it obtained*** (provided to you by upper-level DoD officials, other staff within your installation, out of your own interests, etc.)?

    b. How is it ***applied within your job/your installation*** (what decisions are made and what other factors are important in these decisions)?

    c. What ***types of data are used*** (and for what purpose/routinely/infrequently)?

    i. Estimates of predicted mean temperature/precipitation values (annual/seasonal/monthly/daily).

    ii. Estimates of predicted weather extremes (if so which ones).

d. What is the **value of this information** to you and your installation (high/low)?

e. What **factors** limit your/your installation's use of climate change estimates?

i. Source/s unknown

ii. Accuracy…what level/s desired/needed (is uncertainty stated)

iii. Credibility of source (U.S. vs. foreign model output)

iv. Hard to understand/interpret information provided

v. Information provided not really useful to the installation needs

vi. Others

D. (if they are a user, insert "better" here) If BETTER climate change data/estimates/information were available to you—

1. How would you **use** climate change information (refer to past problems with weather extremes)?

2. What **information** derived from regional climate models would be useful/needed (identify and rank)?
   a. Temperature…mean, maximum, minimum, periods above/below threshold

   b. Precipitation…mean, critical levels being exceeded (within a period of time)

   c. Other weather conditions (e.g., humidity—apparent temperature, wind, sunshine levels, snowfall, ice, hail, lightning, etc.)

   d. Heating/Cooling Degree Days (derived from mean temperature data)

   e. Others

3. What **levels of accuracy/uncertainty** can you live with if you were to use this information?

   a. They will never be perfect, but how would you like the uncertainty characterized/expressed to make it most useful to you?

   b. Are you comfortable with using probabilistic information in your decisions?

4. What other information in predictions would be helpful?

    a. Analogs to similar past periods?

    b. Explanations of the reasons for certainty and uncertainty levels in given predictions (i.e., why is there stronger consensus associated with changes in global mean temperatures than those in a particular region?)

5. What **hindrances** do you see in ever using predicted climate change estimates in decision making efforts at your installation?

    a. Scientific uncertainties?

    b. Lack of process/models to integrate climate change information in a risk analysis or decision process?

    c. No support from others above you (in/out of installation) for using such information (lack of policy/mandate from above).

E. Summary

1. Any final comments or questions about weather or climate issues?

2. Our follow-up…may call you to get clarification on an issue we are not clear on. A report synthesizing our findings from all installations will be sent to all participants.

3. Names of persons who do your job in other installations (who might be approachable to a similar Q & A session).

4. Your email address, telephone number, and address.

5. We appreciate your time and cooperation.

Appendix 2. List of DoD Installations used in this part of the Project

Army:

Aberdeen Proving Ground
Fort Drum
Fort Hood
Fort McCoy
Fort Riley
Fort Stewart
JB Lewis-McChord
Yuma Proving Ground

Air Force:

Tyndall AFB

Marine Corps:

Camp Pendleton

APPENDIX B

List of Scientific/Technical Publications

1. Jin, Z., Zhuang, Q., Wang, J., Archontoulis, S. V., Zobel, Z. and Kotamarthi, V. R. (2017), The combined and separate impacts of climate extremes on the current and future US rainfed maize and soybean production under elevated CO_2. Glob Change Biol. doi:10.1111/gcb.13617
2. Zobel, Z., J. Wang, D. J. Wuebbles, and V. R. Kotamarthi (2017). Evaluations of high-resolution dynamically downscaled ensembles over the contiguous United States. Accepted by Climate Dynamics.
3. Wang, J., Y. Han, M. Stein, V. R. Kotamarthi, and W. K. Huang (2016): Evaluation of dynamical downscaled extreme temperature using a spatially-aggregated generalized extreme value (GEV) model. *Climate Dynamics*. DOI: 10.1007/s00382-016-3000-3
4. Chang, W., Stein, M. L., Wang, J., Kotamarthi, V. R. and Moyer, E. J. (2016). Changes in Spatio-temporal Precipitation Patterns in Changing Climate Conditions. *Journal of Climate*, 29 (23), 8355-8376.
5. Kotamarthi, V. R, Mearns, L, Hayhoe, K., Castro, C., Wuebbles, D. (2016) 'Use of Climate Information for Decision-Making and Impacts Research: State of Our Understanding', Strategic Environmental Research and Development Program.
6. Wang, J and Kotamarthi, V. R (2015).: High-resolution dynamically downscaled projections of precipitation in the mid and late 21st century over North America, Earth's Future, 3: 268–288. doi:10.1002/2015EF000304.
7. Wang, J., F.N.U. Swati, M. Stein and V. R. Kotamarthi (2015). "Model performance in spatiotemporal patterns of precipitation: New methods for identifying value added by a regional climate model." *Journal of Geophysical Research: Atmospheres, 120, 1239-1259,* DOI: 10.1002/2014JD022434
8. Wang, J. and Kotamarthi V. R (2014): Nested regional climate model (NRCM) downscaling in near-surface fields over Contiguous United States, Volume 119, Issue 14, 27 July 2014, Pages: 8778–8797, *J Geophys. Res*, DOI: 10.1002/2014JD021696.
9. Wang, J., and V. Kotamarthi (2013), Assessment of Dynamical Downscaling in Near-Surface Fields with Different Spectral Nudging Approaches Using the Nested Regional Climate Model (NRCM). J. Appl. Meteor. Climatol. 52, 1576–1591.

This page intentionally left blank

Division
Argonne National Laboratory
9700 South Cass Avenue, Bldg. #240
Argonne, IL 60439

www.anl.gov